REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
27-03-2007	Final Report	10 September 2004 - 11-Mar-09

4. TITLE AND SUBTITLE

Advanced robust STAP algorithms and fast performance evaluation techniques based on rare event theory.

5a. CONTRACT NUMBER
FA8655-04-1-3025

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Professor RAJAN Srinivasan

5d. PROJECT NUMBER

5d. TASK NUMBER

5e. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Twente
EL/TN - 9160, Drienerlolaan 5
7500 AE Enschede 7500 AE
The Netherlands

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

EOARD
Unit 4515 BOX 14
APO AE 09421

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Grant 04-3025

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited. (approval given by local Public Affairs Office)

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report results from a contract tasking University of Twente as follows: The Grantee will investigate the development of new algorithms to reduce the Constant False Alarm Rate (CFAR) detection for radar applications. These will occur in three steps. First, the clutter environment will be modeled to include algorithmic non-homogeneities. Second, the effects of mutual coupling between antenna elements in a STAP array will be considered, either through simulated mathematical models incorporated into processing or through practical measurements. Finally, given the vast amounts of data that need to be processed in STAP, fast simulation methods will be investigated and proposed.

15. SUBJECT TERMS
EOARD, radar

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18, NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			GEORGE W YORK, Lt Col, USAF
UNCLAS	UNCLAS	UNCLAS	UL	121	19b. TELEPHONE NUMBER *(Include area code)* +44 (0)1895 616163

Advanced Robust STAP Algorithms and Fast Performance Evaluation Techniques based on Rare Event Theory

Reports 1, 2, 3, & 4

(March 2007)

Authors: Laura Anitori and Rajan Srinivasan

Telecommunication Engineering Group
University of Twente, The Netherlands

Submitted to:
European Office of Aerospace Research and Development (EOARD)

Award No. FA8655-04-1-3025

(PI: Prof. Rajan Srinivasan)

Program Manager: Dr. George York, Lt. Col. (USAF)

EOARD, London, UK

In collaboration with:

Dr. Muralidhar Rangaswamy

Senior Electronics Engineer
Air Force Research Laboratory/SNHE

Acknowledgement

We wish to acknowledge with appreciation the help and support given to us by Col. Mike Milligan (formerly of EOARD), Dr. Arje Nachman (AFOSR), Col. George York (EOARD), and our direct scientific collaborator Dr. Murali Rangaswamy (AFRL Hanscom).

Rajan Srinivasan
March 2007

Summary

This report describes the development of adaptive importance sampling techniques for estimating false alarm probabilities of detectors that use space-time adaptive processing (STAP) algorithms. The first paper to appear on this topic is [9]; it lays the groundwork for developing powerful estimation algorithms, based on the work in fast simulation carried out by the author of [5]. The report also presents some new detection algorithms that are intended to be robust under various conditions.

Fast simulation using importance sampling methods have been notably successful in the study of conventional constant false alarm rate radar detectors, and in several other applications. The principal objectives here are to examine the viability of using importance sampling methods for STAP detection, develop these methods into powerful analysis and design algorithms, and use them for synthesizing novel detection structures.

Several STAP detectors have been investigated from the standpoint of applying importance sampling to characterize their performances. Various biasing techniques have also been devised and implemented, resulting in significant speed-ups in performance evaluation compared to conventional Monte Carlo methods. The important problem of detector threshold determination has been addressed and solved by fast simulation.

Robust variants such as the envelope-law and geometric-mean detectors for STAP processing have been suggested, their CFAR property established, and performance thoroughly evaluated using IS techniques. It has been shown that their detection performances are decidedly better than those of their conventional square-law counterparts when training data are contaminated by interferers, while maintaining almost equal detection performances under homogeneous conditions. The work reported here paves the way to development of more advanced estimation techniques that can facilitate design of powerful and robust detection algorithms designed to counter hostile and heterogenous clutter environments.

During the pendency of the contract, which started in late 2004, the PI and his associates have presented their results at various conferences, in the UK, USA, Europe, and China. One journal paper[1] is due to appear in the IEEE

[1] A pre-print manuscript copy is attached at the end of this report. Conference presentations have not been included here.

Transactions on AES in January 2007 and two more journal manuscripts are to be shortly submitted for review. Owing to unavoidable circumstances the contract has had to be terminated prematurely. Given the satisfactory results obtained in this short time frame, it is hoped that more work will be carried out along the same lines in the present research area.

Contents

List of Figures

Chapter 1

Importance sampling for STAP detection analysis and design - the AMF

1.1 Introduction

Estimation of false alarm probabilities of detection algorithms that employ space-time processing is examined here using forced Monte Carlo or importance sampling simulation (IS). Space-time adaptive processing (STAP) algorithms are of much importance for radar detection. They are notoriously intensive from a computational point of view, with the more advanced (and robust) ones being also analytically difficult to quantize, [3]. Therefore it is appropriate to attempt to develop fast simulation methods that could be used in their analysis and design.

In this chapter we use lessons learnt from developing IS techniques for characterizing conventional constant false alarm rate (CFAR) detectors, [5], and describe an experiment in applying them to STAP detection. The starting (and ending) point of this unpretentious effort is the celebrated adaptive matched filter (AMF) derived in [1] and which represents the arrayed version of the workhorse cell averaging CFAR detector for conventional radar signal processing algorithms. The false alarm probability (FAP) performance of the AMF detector is known in integral form and can be numerically computed to any desired accuracy. Thus it forms a suitable basis for validating our simulation experiments. Two specific IS methods (described in the sequel) are presented and the better (and also easier) one is implemented. On a general note, IS is the chief simulation methodology for rare event estimation. It is an enduring method that has distinguished itself in several areas of science and engineering. Briefly, IS works by biasing original probability distributions in ways that accelerate the occurrences of rare events, conducting simulations with these new

distributions, and then compensating the obtained results for the changes made. The principal consequence of this procedure is that unbiased probability estimates with low variances are obtained quickly. The main task in IS therefore is determination of good simulation distributions for an application, either as a one-shot feat or adaptively. Simulations performed using such distributions can provide enormous speed-ups if they are chosen with due care and mathematical precision. Indeed, if applied successfully, simulation lengths needed to estimate very low probabilities become (only) weakly dependent on the actual probabilities. It is thus possible to evaluate any probability in reasonable amounts of simulation time. There have been more recent attempts in the literature, for example [13], [14], to apply IS for FAP estimation of conventional CFAR detectors with varying degrees of success.

During the conduct of simulations reported herein, some issues concerning the adaptive IS algorithms used arise, and these are discussed briefly. More investigation is required into them. However, the positive outcome of the methods used is that excellent match with numerical results is obtained. The succeeding sections provide a short statement of the AMF algorithm, how IS biasing can be performed to hasten false alarm events, description of the so called g-method which is a conditional IS technique developed originally for studying sums of random variables ([6]), the fast algorithms used, how *inverse IS* can be used to estimate (and therefore design) detector thresholds, simulation results, and a concluding discussion. The work carried out in this chapter has appeared in [9].

1.2 The AMF detector

In a radar system consisting of a linear array of N_s antenna elements, a burst of N_t pulses is transmitted and each element receives as many return samples in any one range gate. The $N_sN_t = N$ samples are complex (because of I and Q channel processing) and are referred to as the primary data. They may contain a target and represent the range gate to be tested. The samples are arranged in an $N \times 1$ column vector and denoted as $\mathbf{x}$. The target return is modelled as consisting of a known direction vector $\mathbf{s}$ with an unknown complex amplitude in addition to clutter, interference, and noise. There are L other N-length complex vectors, called the training data, obtained from as many nearby range gates and assumed to be free of target signal. These are denoted as $\mathbf{x}(l)$, $l = 1, \ldots, L$. It is assumed from now on that the training data is free of other targets or contamination[1]. The primary and secondary data vectors are assumed to be jointly independent and complex Gaussian, sharing the $N \times N$ covariance matrix $\mathbf{R} = E\{\mathbf{X}\mathbf{X}^\dagger\}$, where the superscript $\dagger$ denotes complex conjugate transpose.

[1] In the second chapter, detection performance is carried out for training data contaminated by interfering targets.

Under these assumptions the AMF detection test, as obtained in [1], is given by

$$\frac{|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}|^2}{\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{s}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{1.1}$$

where

$$\widehat{\mathbf{R}} \equiv \frac{1}{L}\sum\nolimits_{l=1}^{L} \mathbf{x}(l)\mathbf{x}(l)^\dagger$$

is the estimated covariance matrix of $\mathbf{x}$ based on the training data (also referred to as sample matrix), and η is a threshold used to set the FAP at some desired level. This test has the CFAR property. The FAP α of the test is known to be given by

$$\alpha = \frac{L!}{(L-N+1)!(N-2)!}\int_0^1 \frac{x^{L-N+1}(1-x)^{N-2}}{(1+\eta\, x/L)^{L-N+1}}\, dx \tag{1.2}$$

which can be used to numerically determine the threshold setting for a desired FAP. As shown in [1], the test in (1.1) can be rewritten as

$$\begin{aligned}
|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}|^2 \underset{H_0}{\overset{H_1}{\gtrless}} & \ \eta\, \mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{s} \\
= & \ \eta\, \mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\widehat{\mathbf{R}}\widehat{\mathbf{R}}^{-1}\mathbf{s} \\
= & \ \eta\, \mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\frac{1}{L}\sum\nolimits_{l=1}^{L} \mathbf{x}(l)\mathbf{x}(l)^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{s} \\
= & \ \frac{\eta}{L}\sum\nolimits_{l=1}^{L} \mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}(l)\mathbf{x}(l)^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{s} \\
= & \ \frac{\eta}{L}\sum\nolimits_{l=1}^{L} |\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}(l)|^2
\end{aligned} \tag{1.3}$$

This is in the form of a vector (or, array) version of the usual CA-CFAR test. The LHS is a square law detector, being the output of a matched filter (matched to the direction $\mathbf{s}$ in which the array is steered) for incoherent detection using the so-called sample matrix inversion (SMI) beamformer weights $\widehat{\mathbf{R}}^{-1}\mathbf{s}$. The RHS represents a cell averaging term. Further details on these issues can be found in the references mentioned above.

1.3 False alarm probability estimation using IS

Two methods to quickly estimate FAPs are two-dimensional (2-d) biasing and the conditional g-method procedure, described in this section.

1.3.1 2-d biasing

To estimate FAP using IS, we make the following observations. Suppose each complex sample of a training vector is scaled by a real number $\theta^{1/2}$. This has the effect of scaling the covariance matrix estimate $\widehat{\mathbf{R}}$ by θ. Therefore, as far as the covariance estimate is concerned, both sides of the test in (1.3) remain

unaffected by the scaling. However, each training vector being scaled by $\theta^{1/2}$ results in a scaling of the RHS by θ. Hence choosing θ less than unity will have the effect of compressing the density function of the random threshold of the test. Further, a scaling of each complex component of the primary vector by a real $a^{1/2}$ will achieve a scaling of the LHS of the test by a. Thus, choosing a larger and θ smaller than unity will achieve an increase in the frequency of occurrence of a false alarm event during simulation. The IS optimization problem will be a two-parameter one.

The (unbiased) IS estimator, using (1.1), can be expressed as

$$\widehat{\alpha} \;=\; \frac{1}{K}\sum\nolimits_1^K 1(|\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}|^2 \;>\; \eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s})\,W(\mathbf{X},\mathbf{X}_L;a,\theta); \quad\sim\quad f_\star \quad (1.4)$$

where the notation $\sim f_\star$ means that all random variables are drawn from biased distributions, and $\mathbf{X}_L \equiv (\mathbf{X}(1),\ldots,\mathbf{X}(L))^T$ with K denoting length of the IS simulation. In setting up their joint densities, we use the fact that the FAP of the AMF has the CFAR property and is independent of the true covariance matrix $\mathbf{R}$. This is true under the assumption of Gaussian distributions for the data. In such a case, the simulation of the AMF test can be carried out for data possessing a diagonal covariance matrix $\mathbf{I}$, denoting the $N\times N$ identity matrix. Therefore, primary and training data can be generated as complex vectors with independent components. The unbiased joint densities are

$$f(\mathbf{x}) = \frac{e^{-\mathbf{x}^\dagger\mathbf{x}}}{\pi^N} \;\text{ and }\; f(\mathbf{x}_L) = \frac{e^{-\sum_1^L \mathbf{x}(l)^\dagger\mathbf{x}(l)}}{\pi^{LN}}$$

so that

$$f(\mathbf{x},\mathbf{x}_L) = \frac{e^{-\mathbf{x}^\dagger\mathbf{x}-\sum_1^L \mathbf{x}(l)^\dagger\mathbf{x}(l)}}{\pi^{(L+1)N}}$$

With scaling performed as described above, the biased joint density takes the form

$$f_\star(\mathbf{x},\mathbf{x}_L) = \frac{e^{-\frac{1}{a}\mathbf{x}^\dagger\mathbf{x}-\frac{1}{\theta}\sum_1^L \mathbf{x}(l)^\dagger\mathbf{x}(l)}}{\pi^{(L+1)N}a^N\theta^{LN}}$$

resulting in the weighting function

$$\begin{aligned} W(\mathbf{X},\mathbf{X}_L;a,\theta) \;&\triangleq\; \frac{f(\mathbf{x},\mathbf{x}_L)}{f_\star(\mathbf{x},\mathbf{x}_L)} \\ &=\; Ca^N\theta^{LN}e^{A/a}e^{B/\theta} \end{aligned} \quad (1.5)$$

where

$$A \equiv \mathbf{x}^\dagger\mathbf{x}, \quad B \equiv \sum\nolimits_1^L \mathbf{x}(l)^\dagger\mathbf{x}(l), \text{ and } C \equiv e^{-(A+B)}$$

The variance of the IS estimator $\widehat{\alpha}$ can be expressed as

$$\operatorname{var}\widehat{\alpha} = \frac{1}{K}[I(\boldsymbol{\nu}) - \alpha^2] \quad (1.6)$$

where $\boldsymbol{\nu}$ is the vector biasing parameter $(a,\theta)^T \in [1,\infty)\times(0,1]$. Denoting by $\mathcal{A}$ the false alarm event in (1.4), the quantity I is given by

$$\begin{aligned} I(\boldsymbol{\nu}) &= E_\star\{1(\mathcal{A})W^2(\mathbf{X},\mathbf{X}_L;\boldsymbol{\nu})\} \\ &= E\{1(\mathcal{A})W(\mathbf{X},\mathbf{X}_L;\boldsymbol{\nu})\} \end{aligned} \tag{1.7}$$

where the expectation $E_\star$ proceeds over biased distributions. Minimization of $\operatorname{var}\widehat{\alpha}$ with respect to the biasing parameters is equivalent to minimization of I and is described in the Appendix. Although not implemented here, this description has been included since it is foreseen that such a method could be useful in situations wherein the g-method might be difficult to apply.

1.3.2 The g-method estimator

This method exploits knowledge of underlying distributions more effectively, yielding a more powerful estimator. Additional advantages are that only a scalar parameter optimization problem needs to be tackled and the *inverse IS* problem (for threshold optimization or selection) can be easily solved. The FAP can be written as

$$\begin{aligned} \alpha &= P(|\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}|^2 > \eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}|H_0) \\ &= E\{P(|\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}|^2 > \eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}|\mathbf{X}_L,H_0)\} \\ &\triangleq E\{g(\mathbf{X}_L)\} \end{aligned} \tag{1.8}$$

Note that the conditioning in the second step above is equivalent to the condition that a covariance matrix estimate is given. We proceed to estimate α using the form in the third step above.

With the condition in mind it is easy to show, assuming that $\mathbf{X}$ is rotationally invariant and Gaussian, that the random variable $\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X} \triangleq \mathbf{w}^\dagger\mathbf{X}$ is distributed as $\mathcal{CN}_1(0,\mathbf{w}^\dagger\mathbf{R}\mathbf{w})$ with independent real and imaginary components, and the weight vector $\mathbf{w} = \widehat{\mathbf{R}}^{-1}\mathbf{s}$. The random variable $Y \triangleq |\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}|^2$ therefore is exponential and has density function

$$f(y|\mathbf{X}_L,H_0) = \frac{e^{-y/\mathbf{w}^\dagger\mathbf{R}\mathbf{w}}}{\mathbf{w}^\dagger\mathbf{R}\mathbf{w}}, \quad y \geq 0$$

Therefore

$$\begin{aligned} g(\mathbf{X}_L) &= P(Y \geq \eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}|\mathbf{X}_L,H_0) \\ &= e^{-\eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}/\mathbf{w}^\dagger\mathbf{R}\mathbf{w}} \end{aligned}$$

Note that if $\widehat{\mathbf{R}} = \mathbf{R}$, then $g(\mathbf{X}_L) = e^{-\eta}$ and this is the FAP of the AMF when the covariance matrix is known. As discussed before, we are simulating with homogeneous data possessing an identity covariance matrix, that is, with $\mathbf{R} = \mathbf{I}$. The g-method IS estimator then takes the form

$$\widehat{\alpha}_g = \frac{1}{K}\sum\nolimits_1^K g(\mathbf{X}_L)W(\mathbf{X}_L;\theta)$$

$$= \quad \frac{1}{K}\sum_1^K e^{-\eta\, D} W(\mathbf{X}_L;\theta); \quad \sim f_\star \tag{1.9}$$

where

$$\begin{aligned} D &\equiv \frac{\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{s}}{|\mathbf{w}|^2} \\ &= \frac{\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{s}}{\mathbf{s}^\dagger (\widehat{\mathbf{R}}^{-1})^2\mathbf{s}} \end{aligned} \tag{1.10}$$

Choosing the (single) biasing parameter $\theta < 1$ thus produces a decrease in D, thereby causing a higher frequency of occurrence of the false alarm event or, more appropriately in this case, a larger value of the g-function. Note that use of the g-method obviates the need to bias primary data vectors. Determination of a good value of θ proceeds as before. The weighting function is simply

$$W(\mathbf{x}_L;\theta) = \theta^{LN} e^{-(1-1/\theta)B} \tag{1.11}$$

which can be deduced from (1.5) by setting $a = 1$. The scaling θ is optimized by

$$\theta_{m+1} = \theta_m - \delta_\theta \frac{\widehat{I}'_g(\theta_m)}{\widehat{I}''_g(\theta_m)} \tag{1.12}$$

which is just a one-dimensional version of (1.16). Estimates of the I-function and its derivatives are given by

$$\begin{aligned} \widehat{I}_g(\theta) &= \frac{1}{K}\sum_1^K g^2(\mathbf{X}_L) W^2(\mathbf{X}_L;\theta); \quad \sim f_\star \\ \widehat{I}'_g(\theta) &= \frac{1}{K}\sum_1^K g^2(\mathbf{X}_L) W(\mathbf{X}_L;\theta) W_\theta(\mathbf{X}_L;\theta); \quad \sim f_\star \\ \widehat{I}''_g(\theta) &= \frac{1}{K}\sum_1^K g^2(\mathbf{X}_L) W(\mathbf{X}_L;\theta) W_{\theta\theta}(\mathbf{X}_L;\theta); \quad \sim f_\star \end{aligned}$$

See appendix 1.5 for definition of the above quantities.

Simulation gain

A useful measure of the effectiveness of any IS procedure is the simulation gain Γ. It is the ratio of simulation lengths required by conventional MC and IS estimators to achieve the same error variance. Setting the variance in (1.6) equal to $(\alpha - \alpha^2)/k$ (being the MC variance) where k denotes the length required by the MC estimator, yields the gain

$$\Gamma = \frac{\alpha - \alpha^2}{I(\boldsymbol{\nu}) - \alpha^2}$$

While the simulation gain is useful in learning how much faster than MC an IS technique is in terms of simulation length, it also serves the purpose of comparing

different IS estimators. In actual simulations, an estimate of Γ is made by using the estimates for α and I. The g-method estimator has simulation gain given by

$$\Gamma_g = \frac{\alpha - \alpha^2}{I_g(\boldsymbol{\theta}) - \alpha^2}$$

where $I_g = E_\star\{g^2(\mathbf{X}_L)W^2(\mathbf{X}_L;\theta)\}$, and it can be estimated during simulation. It always has a smaller variance and consequently larger gain than the IS estimator discussed in the previous section. Indeed, without IS ($W = 1$), $I_g = E\{g^2(\mathbf{X}_L)\} < E\{g(\mathbf{X}_L)\} = \alpha$. That $I_g < I$ with IS was proved in [5] for conventional CFAR detectors. The proof in the case of the detectors considered here is similar, and will be omitted.

1.3.3 Inverse IS and threshold determination

The inverse problem, namely that of finding by fast simulation the value of detector threshold η satisfying a prescribed FAP, is readily solved using the g-method estimator. This is done by minimizing the stochastic objective function

$$J(\eta) = [\widehat{\alpha}_g(\eta) - \alpha_o]^2$$

where α_o is a desired FAP. An example is shown in Figure 1.1. It is clear that all detection algorithms that involve a threshold crossing will possess objective functions that have the general behaviour shown, assuming of course that the FAP estimate is a decreasing function of its argument η_g. Minimization of J with respect to η is carried out by the algorithm

$$\eta_{m+1} = \eta_m + \delta_\eta \frac{\alpha_o - \widehat{\alpha}_g(\eta_m)}{\widehat{\alpha'_g}(\eta_m)}, \quad m = 1, 2, \ldots \tag{1.13}$$

where δ_η is a step-size parameter and the derivative estimator in the denominator is given by

$$\widehat{\alpha'_g}(\eta_m) = -\frac{1}{K}\sum\nolimits_1^K D\, e^{-\eta D} W(\mathbf{X}_L;\theta); \quad \sim f_\star \tag{1.14}$$

with the prime indicating derivative with respect to η. Note in passing that this derivative estimator actually estimates (negative of) the probability density function of the AMF statistic on the left hand side of (1.1) under H_0.

1.4 Numerical results

The FAP α obtained by direct numerical integration of (1.2) is shown in Figure 1.2 and is used for comparing IS results, which are displayed in the remaining figures. The AMF detector consists of $L = 704$ trainng vectors each of length $N = 352$. Shown in Figure 1.3 is one instance of adaptive IS estimation of FAP for a (known) threshold of $\eta = 56.50432$. Figures 1.4 and 1.5 depict results from

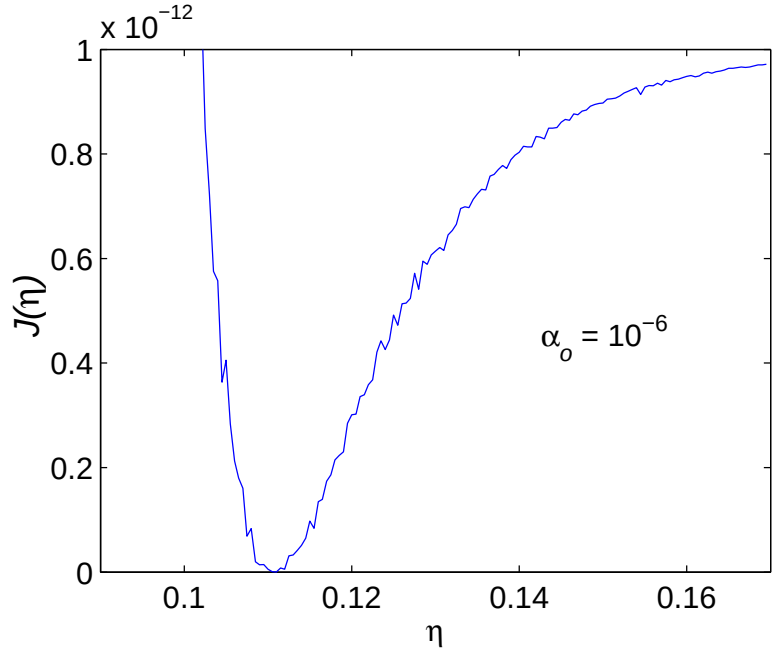

Figure 1.1: An objective fucntion.

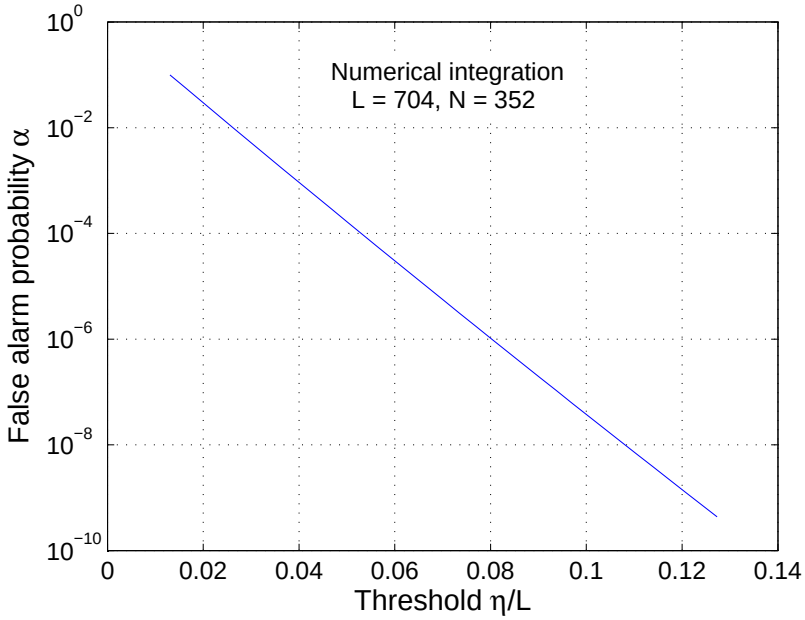

Figure 1.2: Numerically computed FAP of the AMF detector.

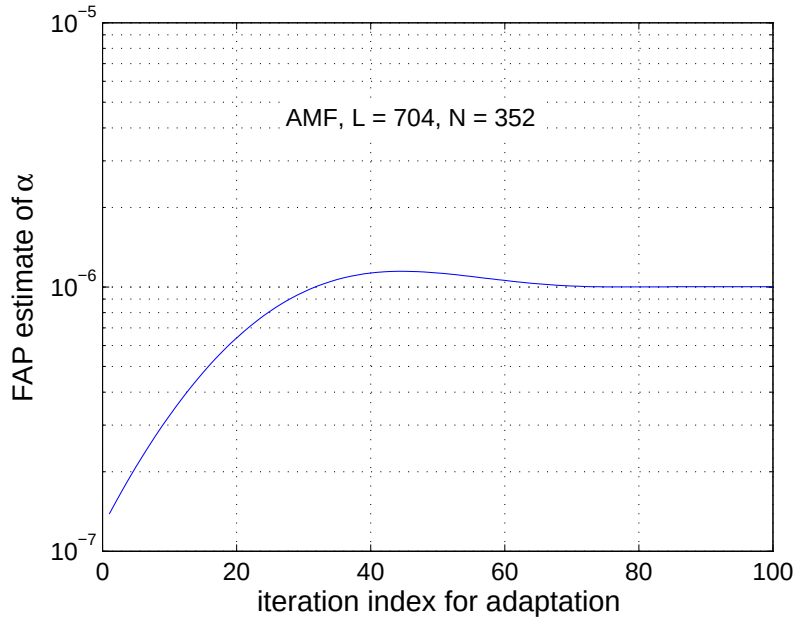

Figure 1.3: Convergence of FAP using adaptive IS algorithms.

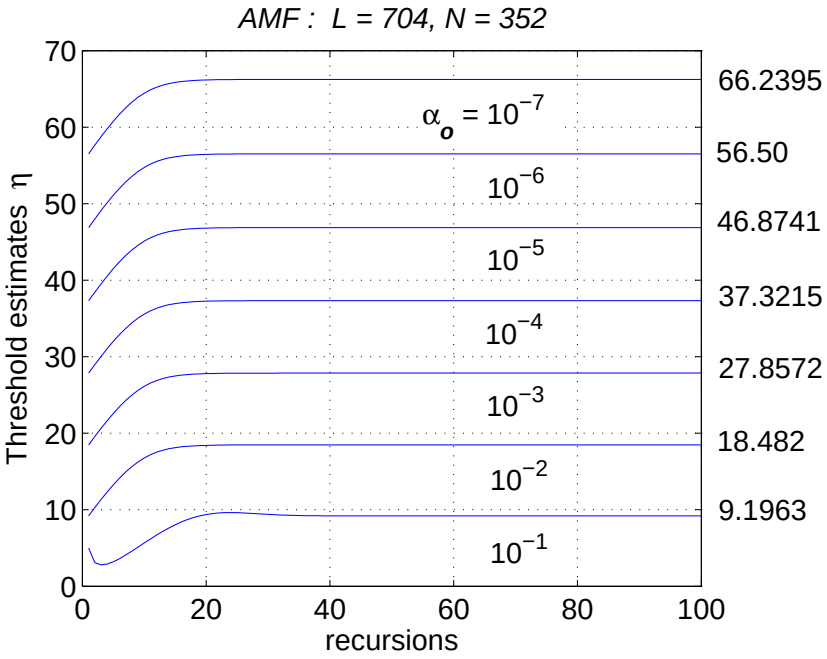

Figure 1.4: Threshold optimization for AMF detector using inverse IS.

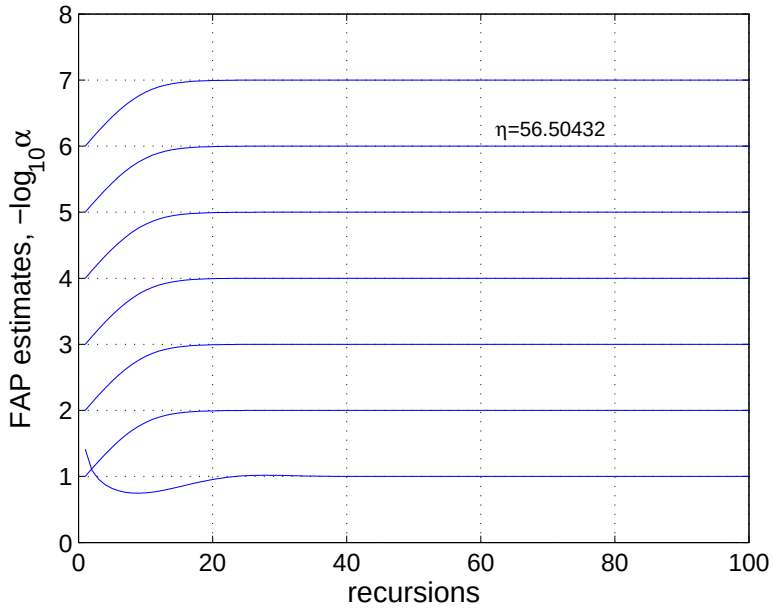

Figure 1.5: FAP estimates resulting from threshold optimization algorithms.

implementing the inverse IS algorithms. These are estimated threshold settings and associated FAPs respectively. It is evident that match with the results in Figure 1.2 is excellent and this has been numerically confirmed.

Optimum biasing parameters are shown in Figure 1.6 and simulation gains obtained in Figure 1.7. Shown in Figure 1.8 is the gain as function of FAP. The number of trials K needed to provide an accuracy of $\pm 10\%$ with 95% confidence is shown in Figure 1.9.

1.4.1 Discussion

The IS simulation results obtained here appear deceptively smooth and certainly beg an obvious question. Indeed, an artifice has been employed here to generate them. It was used by the first author in previous work on capacity estimation of MIMO channels and elsewhere, and found to be extremely useful. In conducting rare-event simulations of systems that involve signal processing operations that are mathematically complex, there are two principal issues that contribute to simulation time. These have to be dealt with effectively. The first issue concerns the rare event itself whose probability is being sought, and this can of course be handled by suitable IS techniques. The second concerns the computational intensity that accompanies the signal processing. The two are not unrelated.

In the case of STAP detectors, the chief processing burden is from inversion of the sample matrix. It is a daunting task to conduct conventional Monte

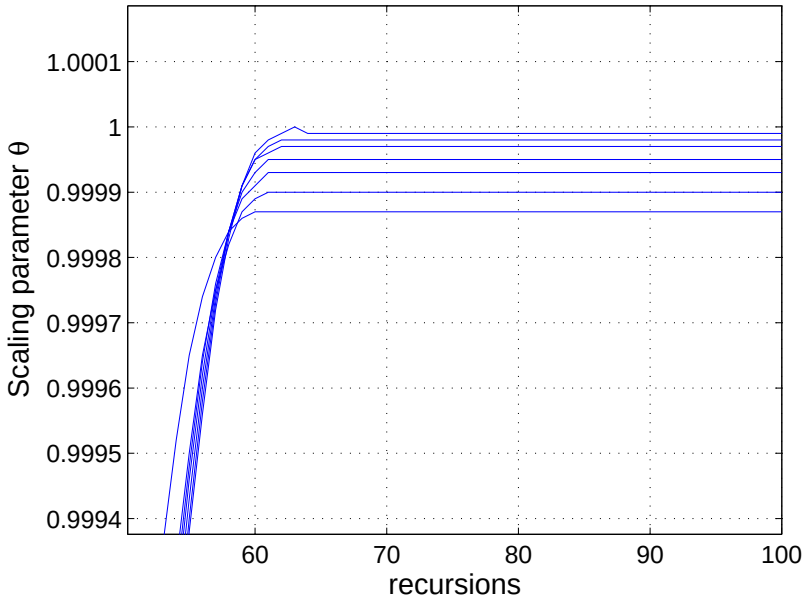

Figure 1.6: Optimum biasing (scaling) parameters for IS.

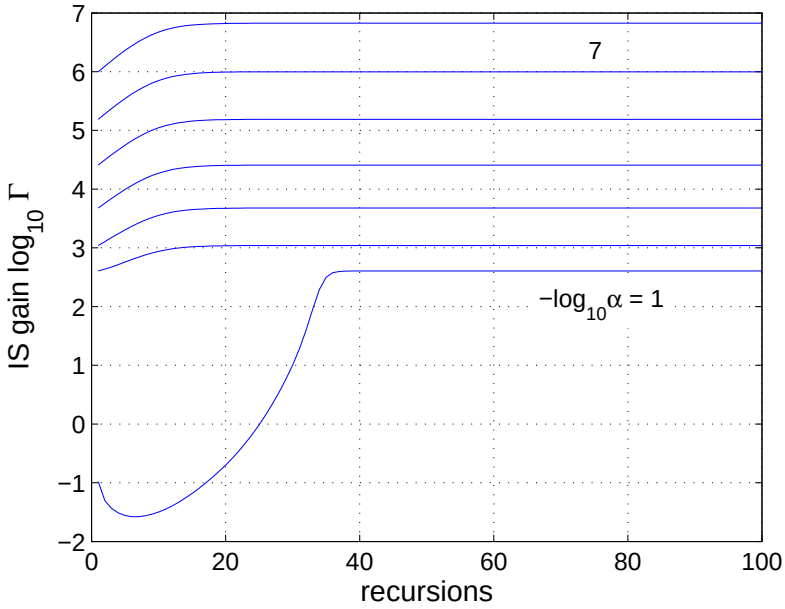

Figure 1.7: Estimated simulation gains.

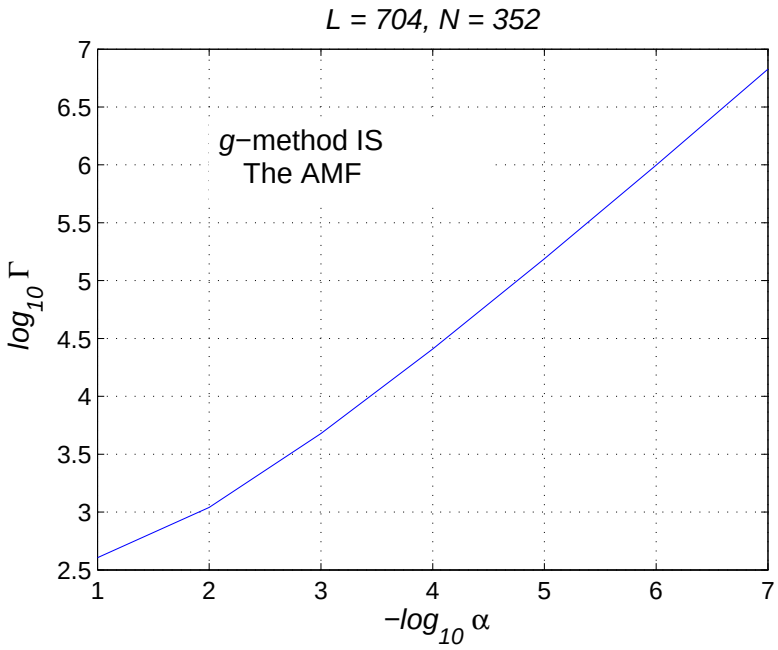

Figure 1.8: Simulation gains.

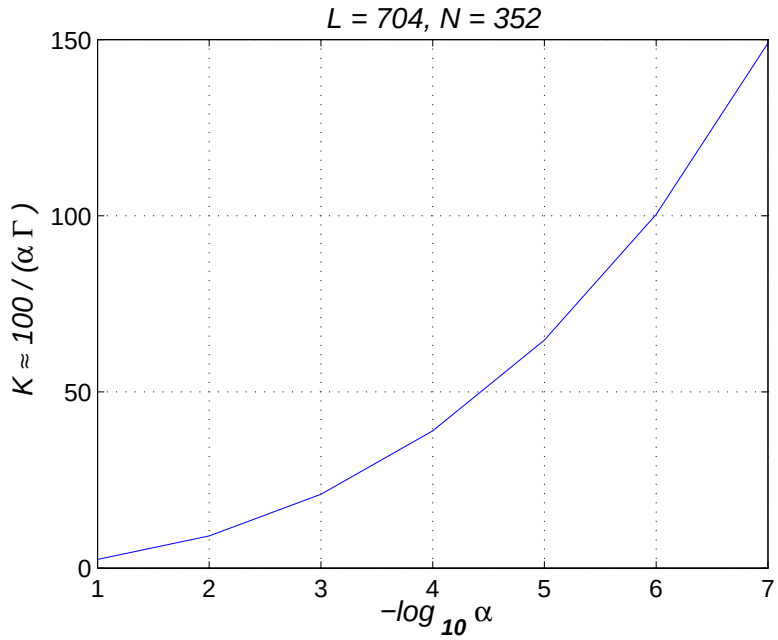

Figure 1.9: Simulation trials K required.

Carlo simulations that involve several millions of trials to estimate low FAPs, with as many matrix inversions. Assuming that a good IS scheme can reduce the number of trials to, say, only a few thousands would still be computationally demanding (a case in point being the three thousand 352×352 matrices that were inverted here). This is the point at which IS departs from conventional Monte Carlo in a subtle but important manner. It is almost totally useless to run the same random variables through a system in a straight MC simulation. With IS however, much can be learnt by repeatedly using the (same) random variables. In fact, this is one of the powerful features that adaptive IS (and inverse IS) can embed into complex system simulation.

But how does an IS scheme become effective in the first place? Assume that we have a biasing scheme that promises to be effective once the parameters of the biasing distributions have been optimized. For large systems (in the sense of number of random inputs involved), running truly randomized IS algorithms adaptively could become demanding as pointed out above. If system performance can be characterized in terms of certain random 'metrics' (we use the word with a slight abuse of terminology), then these metrics can be pre-computed for a given set of input variables, and used repeatedly (which, in complex systems such as STAP detectors eases the computational burden) in adaptive biasing optimization algorithms. These latter algorithms themselves usually require no more than 100 iterations and can be extremely fast. Resulting IS simulation gains can be simultaneously estimated and these tell us whether we need more or less pre-generated variables to achieve certain accuracies. Adjusting this latter number, biasing and system parameter optimization (inverse IS) algorithms can be run, once. Thus there is an initial stage of at most a few steps during which gains are estimated based on pre-computed metrics and the number of these metrics is adjusted.

All this is not as complicated as it appears. Turning attention to (1.9), (1.11), and (1.14), the only two random quantities (or metrics) that are needed to estimate the FAP and associated detection threshold are B and D. This is for the g-method. For 2-d biasing the only additional quantity required is the norm A of the primary vector, defined just after (1.5) and this adds almost nothing to the computation. It turned out that generating $K = 3000$ random instances of B and D was certainly an overkill. If one looks at Figure 1.7, the gain provided by IS for estimating $\alpha = 10^{-6}$ is about 10^6. From usual asymptotic normality arguments, [5], it follows that about 100 optimally biased trials are sufficient to guarantee an absolute estimation accuracy not exceeding 10% with 95% confidence. For 2-d biasing, the simulation gain will be somewhat lower but the essential advantages of the method above remain. That is, handling a few hundred inversions (once) is not at all a tall order. This method can produce such an avalanche of results that it is tempting to think of it (with a slight stretch of imagination) as a 'turbo-IS'. The above ideas certainly need quantification but it is beyond the remit of this short report to delve deeper.

An interesting observation comes from Figure 1.6, which shows that the biasing parameter is very close to unity and has a small spread despite the range of FAPs considered. The implication is that the (one-sided) density of

the metric D has small variance, presumably owing to the choice of L and N. Smaller values of these constants would probably lead to larger spread of biasing parameter. In the actual adaptations, a small value of the step-size parameter was used to ensure gradual but safe convergence. This explains their apparently slow nature as seen in the figures. While configuring results for a suite of system parameters, only the first adaptation need be somewhat long; subsequent adaptations can be much shorter as they pick up starting or initial values from the previous one.

1.5 Conclusions

We have made a small inroad into the use of adaptive IS algorithms to characterize a STAP detector. The AMF was used as example and results have been very good. The chief reasons for this are that we were able to invoke the g-method and inverse IS, find a suitable biasing strategy that could be easily optimized adaptively, and find a way around the difficult task of inverting large matrices several times (as described above). The hope is that applications to other STAP configurations, such as normalized AMF and those that handle non-homogenous clutter, will also meet with success. But this remains to be seen as we are certainly not in position to predict what subtleties these other detection algorithms can throw up. It is clear that IS is still in its infancy, especially insofar as use for characterizing modern detection algorithms is concerned. The simulation experiments conducted here have raised questions that need to be answered subsequently.

Appendix

Adaptive algorithms for 2-d biasing The I-function is estimated as

$$\widehat{I}(\boldsymbol{\nu}) = \frac{1}{K}\sum\nolimits_1^K 1(\mathcal{A})\, W^2(\mathbf{X}, \mathbf{X}_L; \boldsymbol{\nu}); \quad \sim f_\star \tag{1.15}$$

and its minimization can be carried out using the 2-dimensional adaptive algorithm

$$\boldsymbol{\nu}_{m+1} = \boldsymbol{\nu}_m - \delta \widehat{\mathbf{J}}_m^{-1} \widehat{\nabla} I(\boldsymbol{\nu}_m) \tag{1.16}$$

Here, δ is a step-size parameter used to control convergence, and m is the index of recursion. This is a stochastic Newton recursion. It achieves minimization of $\widehat{I}$ by estimating a solution of

$$\widehat{\nabla} I(\boldsymbol{\nu}) \equiv (\widehat{I}_a \quad \widehat{I}_\theta)^T = \mathbf{0}$$

where $I_a \triangleq \partial I(\boldsymbol{\nu})/\partial a$ and $I_\theta \triangleq \partial I(\boldsymbol{\nu})/\partial \theta$. The estimate of the Jacobian $\mathbf{J}$ (which is the Hessian matrix of I) is given by

$$\widehat{\mathbf{J}} = \begin{pmatrix} \widehat{I}_{aa} & \widehat{I}_{a\theta} \\ \widehat{I}_{a\theta} & \widehat{I}_{\theta\theta} \end{pmatrix}$$

where $I_{xy} \equiv \partial I_x/\partial y$. It is straightforward to show that the various I-functions defined above can be obtained by the notational equations

$$\begin{aligned} I_x &= E_\star\{1(\mathcal{A})WW_x\} \\ I_{xx} &= E_\star\{1(\mathcal{A})WW_{xx}\} \\ I_{xy} &= E_\star\{1(\mathcal{A})WW_{xy}\} \end{aligned}$$

with various derivatives of the weighting function calculated as

$$\begin{aligned} W_a &\equiv \frac{\partial W}{\partial a} \\ &= \Big(N - \frac{A}{a}\Big)\frac{W}{a} \\ W_\theta &\equiv \frac{\partial W}{\partial \theta} \\ &= \Big(LN - \frac{B}{\theta}\Big)\frac{W}{\theta} \\ W_{aa} &\equiv \frac{\partial^2 W}{\partial a^2} \\ &= \Big[\Big(N - \frac{2A}{a}\Big)(N-1) + \frac{A^2}{a^2}\Big]\frac{W}{a^2} \\ W_{\theta\theta} &\equiv \frac{\partial^2 W}{\partial \theta^2} \\ &= \Big[\Big(LN - \frac{2B}{\theta}\Big)(LN-1) + \frac{B^2}{\theta^2}\Big]\frac{W}{\theta^2} \\ W_{a\theta} &\equiv \frac{\partial^2 W}{\partial a \partial \theta} \\ &= \Big(LN - \frac{B}{\theta}\Big)\Big(N - \frac{A}{a}\Big)\frac{W}{a\theta} \end{aligned}$$

and they can be estimated as in (1.15). The FAP estimator in (1.4) and the adaptive biasing algorithm of (1.16) are implemented simultaneously.

Chapter 2

Geometric-Mean and Envelope-AMF STAP Detectors

2.1 Introduction

In the previous chapter we considered the FAP and threshold estimation of AMF detector by using IS and inverse IS respectively. Beside FAP, another performance measure of a detector is the detection probability, P_D. The higher the detection probability for certain SNR, the better the detector. For the AMF detector, the expression for detection probability in homogeneous Gaussian background is known (see [1]). If the expression for P_D of a detector is unknown then it is usually estimated by using standard MC simulation. Analytical expression for detection probability of AMF in the presence of interfering targets is unknown. In this chapter, the performance of AMF for interfering targets will be simulated to evaluate its performance.

As noted in the previous chapter, the AMF detector is similar to the cell-averaging (CA) detector (see (1.3)). The CA detector take the sums of magnitude square of the training data. In the scalar case (only 1 antenna element), it has been shown (see Chapter 6 in [5]) that CA detector is better than geometric mean (GM) detector in homogeneous background, but GM detector performs well in the presence of interfering targets. It also has been shown in [7] that mean-level CFAR processors including CA preceded by an envelope detector have more robust performance in the presence of Gaussian clutter power transitions and interfering targets as compared to the square law detector.

Following these facts, we propose the GM-STAP detector and envelope-AMF detector (E-AMF), array processing versions of GM and envelope CA detectors respectively. These STAP detectors should have small loss in homogeneous background case and will show more robustness compared to AMF in the presence

of interfering targets.

In this chapter, we will give the expressions for GM-STAP and E-AMF detector tests and also prove their CFAR property. In the performance evaluation section, FAP and threshold estimation of these detectors by using IS will be given, followed by estimation and comparison of the detection probability of the proposed detectors with AMF in the homogeneous case and in the presence of interfering targets.

2.2 The GM-STAP detector

It is well known that the scalar GM detector test can be formulated, in general, as

$$U \underset{H_0}{\overset{H_1}{\gtrless}} \zeta \left(\prod_{l=1}^{L} U_l \right)^{1/L}$$

where U is the output of the test cell (primary data), U_l's are the outputs of the reference cells (training data), L is the number of reference cells, and ζ is some multiplier. Following the same formulation, we will introduce GM-STAP detector that can be formulated as

$$|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{x}| \underset{H_0}{\overset{H_1}{\gtrless}} \eta_g \left(\prod_{l=1}^{L} |\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{x}(l)| \right)^{1/L} \tag{2.1}$$

where $\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{x}$ and $\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{x}(l)$ are the outputs of the matched filter with RMB beamformer weight $\widehat{\mathbf{R}}^{-1}\mathbf{s}$ applied to test cell and reference cells respectively. The threshold multiplier for this detector is η_g. The definition of $\widehat{\mathbf{R}}$, $\mathbf{s}$, $\mathbf{x}$, $\mathbf{x}(l)$ and L are the same as in previous chapter.

The GM-STAP detector does geometric mean processing of the matched filter ouputs of reference cells whereas the AMF detector computes an arithmetic average.

2.2.1 Known covariance

In the ideal case when the covariance of the data is known, the estimate $\widehat{\mathbf{R}}$ is replaced by $\mathbf{R}$ in the test of (2.1). Normalizing the test by $\sqrt{\mathbf{s}^\dagger \mathbf{R}^{-1} \mathbf{s}}$ yields

$$v \underset{H_0}{\overset{H_1}{\gtrless}} \eta_g \left(\prod_{l=1}^{L} v_l \right)^{1/L}$$

where V and $\{V_l\}$ are all independent. This is exactly the same situation as for the GM-CFAR detector described in [5]. The FAPs of the GM-STAP and GM-CFAR detectors are therefore the same in this case. Furthermore, when the covariance is known, the GM-STAP detector has (almost trivially) the CFAR property. In any case, such a situation is of no practical interest, for if the covariance were known then one would just implement a coherent matched filter with fixed threshold.

Asymptotic threshold

It is of interest, however, to determine the threshold of the GM detector as the number of training vectors $L \to \infty$. For convenience we consider the square-law version of the test in (2.1). Denoting the random part of the (squared) threshold by T_{GM}, we have

$$T_{\text{GM}}(L) = \left(\prod_{l=1}^{L} |\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{x}(l)|^2 \right)^{1/L}$$

so that

$$\log T_{\text{GM}}(L) = \frac{1}{L} \sum_{l=1}^{L} \log |\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{x}(l)|^2$$

Noting that $\widehat{\mathbf{R}}^{-1} \xrightarrow{p} \mathbf{R}^{-1}$ as $L \to \infty$, then $\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{X}(l) \xrightarrow{D} \mathbf{s}^\dagger \mathbf{R}^{-1} \mathbf{X}(l)$ in the absence of target. Therefore, by the law of large numbers

$$\begin{aligned} \log T_{\text{GM}}(L) &\longrightarrow E\{\log |\mathbf{s}^\dagger \mathbf{R}^{-1} \mathbf{X}(l)|^2\} \qquad \text{(in probability)} \\ &= \frac{1}{\mathbf{s}^\dagger \mathbf{R}^{-1} \mathbf{s}} \int_0^\infty (\log x)\, \exp^{-x/\mathbf{s}^\dagger \mathbf{R}^{-1} \mathbf{s}}\, dx \\ &= -\gamma + \log(\mathbf{s}^\dagger \mathbf{R}^{-1} \mathbf{s}) \end{aligned}$$

because $\mathbf{s}^\dagger \mathbf{R}^{-1} \mathbf{X}(l) \sim \mathcal{CN}_1(0, \mathbf{s}^\dagger \mathbf{R}^{-1} \mathbf{s})$ and where γ is the Euler constant. It follows (by Theorem 2.7 of [8]) that

$$\begin{aligned} T_{\text{GM}}(L) &\longrightarrow \exp^{-\gamma} (\mathbf{s}^\dagger \mathbf{R}^{-1} \mathbf{s}) \qquad \text{(in probability)} \\ &= 0.561459\, (\mathbf{s}^\dagger \mathbf{R}^{-1} \mathbf{s}) \\ &\triangleq T_{\text{GM}}(\infty) \end{aligned}$$

The FAP of the GM-STAP detector in this known covariance matrix case is given by

$$\begin{aligned} \alpha_{\text{GM}}(L \to \infty) &= P\left(|\mathbf{s}^\dagger \mathbf{R}^{-1} \mathbf{X}|^2 \geq \eta_g^2\, T_{\text{GM}}(\infty) \big| H_0 \right) \\ &= \exp^{-0.561459\, \eta_g^2} \end{aligned}$$

Note that the constant 0.561459 is the same asymptotic normalized weight ν_{gm} obtained for the scalar GM-CFAR detector in Table 6.3 of [5]. Hence, the threshold η_g^2 required to provide a FAP of 10^{-6} is $13.81551/0.561459 = 24.6064$, in the asymptotic case.

2.3 The envelope-AMF detector

The AMF test described by equation (1.3) in the previous chapter involves the magnitude square of the matched filter ouputs of primary and training

data. We may call this detector as Square Law - AMF (SL-AMF). By a slight modification to this detector, taking the magnitude of the filter outputs instead of the magnitude squared of the filter outputs, we have the test

$$|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}| \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\eta_e}{L} \sum\nolimits_{l=1}^{L} |\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}(l)| \tag{2.2}$$

where the symbols used here have the same meaning as in the previous section and η_e denotes the threshold multiplier for the envelope detector. We call this the envelope-law AMF (E-AMF) detector.

2.4 CFAR property

In this section an invariance property is established that is of use in constructing STAP detection algorithms with FAPs that do not depend on the data covariance $\mathbf{R}$. Although proof of the proposition given here follows the same line of argument essentially contained in the exposition of the generalized likelihood ratio STAP detector test (Kelly's GLRT) given in [2], it is outlined here for convenience of the reader[1]. Assume, as before, that the primary and training data vectors have the same covariance. Consider the variables

$$G \equiv \mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x} \quad \text{and} \quad G(l) \equiv \mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}(l) \tag{2.3}$$

for $l = 1, \ldots, L$, that are involved in the AMF detection test of (1.3). Using the transformations $\mathbf{u} = \mathbf{R}^{-1/2}\mathbf{s}$, $\mathbf{y} = \mathbf{R}^{-1/2}\mathbf{x}$, and $\mathbf{y}(l) = \mathbf{R}^{-1/2}\mathbf{x}(l)$, leads to

$$G = \mathbf{u}^\dagger \widetilde{\mathbf{R}}^{-1}\mathbf{y} \quad \text{and} \quad G(l) = \mathbf{u}^\dagger \widetilde{\mathbf{R}}^{-1}\mathbf{y}(l) \tag{2.4}$$

where $\widetilde{\mathbf{R}} \equiv \mathbf{R}^{-1/2}\widehat{\mathbf{R}}\mathbf{R}^{-1/2}$. The whitened vectors $\mathbf{Y}$ and $\mathbf{Y}(l)$ are both distributed $\mathcal{CN}_N(0, \mathbf{I})$. It turns out that $\widetilde{\mathbf{R}}$ has the complex Wishart distribution[2] $\mathcal{CW}(L, N; \frac{1}{L}\mathbf{I})$. Further, a unitary transformation $\mathbf{U}$ can be found which rotates the new signal vector $\mathbf{u}$ into an elementary vector $\mathbf{e}$ as

$$d\mathbf{e} = \mathbf{U}^\dagger \mathbf{u}$$

such that $\mathbf{e} = [1, 0, \ldots, 0]^\dagger$ and where $d^2 = \|\mathbf{U}^\dagger \mathbf{u}\| = \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s}$. The first column of $\mathbf{U}$ is the new signal vector $\mathbf{u}$. The remaining columns comprise an orthonormal basis determined, for example, by a Gram-Schmidt procedure. Let $\mathbf{z} = \mathbf{U}^\dagger \mathbf{y}$

[1] It would be helpful for the reader to refer to [2] (see also [1]). We have used several results from this now classic paper and have attempted to maintain the same notation.

[2] When $\mathbf{X} \sim \mathcal{CN}_N(0, \mathbf{R})$, the (Wishart) matrix $\mathbf{W} = \sum_{l=1}^{L} \mathbf{X}\mathbf{X}^\dagger = L\widehat{\mathbf{R}}$ has the complex Wishart distribution $\mathcal{CW}(L, N; \mathbf{R})$ specified by the density

$$f_{\mathbf{W}}(\mathbf{w}) = \begin{cases} \frac{(\det \mathbf{w})^{L-N}}{J(\mathbf{R})} \exp(-\mathrm{tr}(\mathbf{R}^{-1}\mathbf{w})), \text{if } \mathbf{R} \text{ is positive definite} \\ 0, \qquad \text{otherwise} \end{cases} \tag{2.5}$$

where $J(\mathbf{R}) = \pi^{N(N-1)/2} \prod_{n=1}^{N} \Gamma(L - n + 1)(\det \mathbf{R})^L$. The covariance estimate $\widehat{\mathbf{R}}$ is distributed as $\mathcal{CW}(L, N; \frac{1}{L}\mathbf{R})$. If $\mathbf{B}$ is an $N \times N$ nonsingular complex matrix, then $\mathbf{V} = \mathbf{B}^\dagger \mathbf{W}\mathbf{B}$ is distributed as $\mathcal{CW}(L, N; \mathbf{B}^\dagger \mathbf{R}\mathbf{B})$. See [4] for more on Wishart distributions.

and $\mathbf{z}(l) = \mathbf{U}^\dagger \mathbf{y}(l)$. Applying these to (2.4) yields the variables

$$G = \frac{d}{L}\,\mathbf{e}^\dagger \mathcal{S}^{-1}\mathbf{z} \quad \text{and} \quad G(l) = \frac{d}{L}\,\mathbf{e}^\dagger \mathcal{S}^{-1}\mathbf{z}(l) \tag{2.6}$$

where $\mathcal{S} \equiv L\,\mathbf{U}^\dagger \widetilde{\mathbf{R}}\mathbf{U}$. While $\mathbf{Z}$ and $\mathbf{Z}(l)$ are distributed as $\mathcal{CN}_N(0, \mathbf{I})$ and are independent, $\mathcal{S}$ has the distribution $\mathcal{CW}(L, N; \mathbf{I})$. The vectors $\mathbf{z}$ and $\mathbf{z}(l)$ are decomposed as

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_A \\ \mathbf{z}_B \end{bmatrix} \quad \text{and} \quad \mathbf{z}(l) = \begin{bmatrix} \mathbf{z}_A(l) \\ \mathbf{z}_B(l) \end{bmatrix}$$

where the A components are scalar and B components $(N-1)$-vector. Correspondingly, $\mathcal{S}$ is decomposed as

$$\mathcal{S} = \sum\nolimits_{l=1}^{L} \mathbf{z}(l)\mathbf{z}(l)^\dagger = \begin{bmatrix} \mathcal{S}_{AA} & \mathcal{S}_{AB} \\ \mathcal{S}_{BA} & \mathcal{S}_{BB} \end{bmatrix} \tag{2.7}$$

with

$$\mathcal{P} \equiv \mathcal{S}^{-1} = \begin{bmatrix} \mathcal{P}_{AA} & \mathcal{P}_{AB} \\ \mathcal{P}_{BA} & \mathcal{P}_{BB} \end{bmatrix}$$

The entries of $\mathcal{P}$ can be expressed as

$$\begin{aligned} \mathcal{P}_{AA} &= (\mathcal{S}_{AA} - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\mathcal{S}_{BA})^{-1} \\ \mathcal{P}_{BA} &= -\mathcal{S}_{BB}^{-1}\mathcal{S}_{BA}\mathcal{P}_{AA} \\ \mathcal{P}_{AB} &= \mathcal{P}_{BA}^{-1} \\ \mathcal{P}_{BB} &= \mathcal{S}_{BB}^{-1} + \mathcal{P}_{AA}^{-1}\mathcal{P}_{BA}\mathcal{P}_{AB} \end{aligned}$$

as shown in [2] (page 120). Using these definitions and relations in (2.6) gives

$$\begin{aligned} G &= \frac{d}{L}\,\mathbf{e}^\dagger \mathcal{S}^{-1}\mathbf{z} \\ &= \frac{d}{L}\,\mathcal{P}_{AA}(\mathbf{z}_A - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\mathbf{z}_B) \\ &= \frac{d}{L}\,\mathcal{P}_{AA}\,y \end{aligned} \tag{2.8}$$

and

$$\begin{aligned} G(l) &= \frac{d}{L}\,\mathbf{e}^\dagger \mathcal{S}^{-1}\mathbf{z}(l) \\ &= \frac{d}{L}\,\mathcal{P}_{AA}\big(\mathbf{z}_A(l) - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\mathbf{z}_B(l)\big) \end{aligned}$$

$$= \quad \frac{d}{L}\,\mathcal{P}_{AA}\,y(l) \tag{2.9}$$

where

$$\begin{aligned} y &\equiv \mathbf{z}_{A} - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\mathbf{z}_{B} \\ y(l) &\equiv \mathbf{z}_{A}(l) - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\mathbf{z}_{B}(l) \end{aligned} \tag{2.10}$$

Conditioned on the vectors $\mathbf{z}_{B}$ and $\mathbf{z}_{B}(l)$, the random variables Y and $Y(l)$ in (2.10) are (in the absence of target) uncorrelated and Gaussian with zero means and variances that can be calculated[3] as

$$\begin{aligned} E_{B}\{|Y|^{2}\} &= 1 + \mathbf{z}_{B}^{\dagger}\mathcal{S}_{BB}^{-1}\mathbf{z}_{B} \\ &= 1 + \mathbf{z}_{B}^{\dagger}\left(\sum\nolimits_{l=1}^{L}\mathbf{z}_{B}(l)\mathbf{z}_{B}(l)^{\dagger}\right)^{-1}\mathbf{z}_{B} \end{aligned} \tag{2.11}$$

using (2.7) in the second step, and

$$E_{B}\{|Y(l)|^{2}\} = 1 - \mathbf{z}_{B}(l)^{\dagger}\mathcal{S}_{BB}^{-1}\mathbf{z}_{B}(l), \qquad l = 1, \ldots, L \tag{2.12}$$

with E_{B} denoting conditional expectation. Further, the conditional covariance of the variables $Y(l)$ is given by

$$E_{B}\{Y(k)Y(n)^{\star}\} = -\,\mathbf{z}_{B}(n)^{\dagger}\mathcal{S}_{BB}^{-1}\mathbf{z}_{B}(k), \quad k \neq n \tag{2.13}$$

Hence the set of conditionally jointly Gaussian zero mean random variables Y and $\{Y(l)\}_{1}^{L}$ have individual variances and covariances that are functions of the random vectors $\mathbf{z}_{B}$ and $\{\mathbf{z}_{B}(l)\}_{1}^{L}$. The latter are all jointly independent, each being distributed as $\mathcal{CN}_{N-1}(0, \mathbf{I})$. The probability of any event defined on the random variables Y and $\{Y(l)\}_{1}^{L}$ in (2.10) can thus be determined by performing an averaging operation over the distributions of $\mathbf{z}_{B}$ and $\{\mathbf{z}_{B}(l)\}_{1}^{L}$ and this probability will be independent of the data covariance $\mathbf{R}$. This statement is also true for the random variables G and $\{G(l)\}_{1}^{L}$ in (2.8) and (2.9) with the caveat that any constant scaling of these variables should leave the event unchanged. The preceding arguments therefore constitute proof of the following

Proposition 1. *Any STAP detection algorithm that uses only the random variables G and $\{G(l)\}_{1}^{L}$ defined in* (2.3) *for its description such that the algorithm itself is unchanged by arbitrary but equal scaling of all these variables, has a FAP which is independent of the target-free data covariance $\mathbf{R}$.*

Proof. As above. □

[3] See pages 121 and 122 of [2].

The proposition above helps to establish the CFAR property of GM-STAP and E-AMF detectors. We can rewrite GM-STAP test in terms of random variables G and $G(l)$ by applying (2.3) to (2.1) as

$$|G| \underset{H_0}{\overset{H_1}{\gtrless}} \eta_g \left(\prod_{l=1}^{L} |G(l)| \right)^{1/L} \tag{2.14}$$

By applying (2.3) to (2.2), we have an expression for E-AMF in terms of G and $G(l)$

$$|G| \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\eta_e}{L} \sum\nolimits_{l=1}^{L} |G(l)| \tag{2.15}$$

Scaling both variables G and $G(l)$ with arbitrary but equal scaling factor in both above equations will not change the tests. So by using proposition 1, it is clear that GM-STAP and E-AMF have the CFAR property.

2.5 Performance Evaluation

In this section we present estimates of the false alarm and detection probability (P_D) of the proposed detectors for several cases. The FAP of proposed detectors cannot be derived analytically so that IS will be used to do the estimation. The inverse IS will be executed to estimate the threshold of proposed detector. The threshold then will be applied to evaluate the detection probability of the detectors. As neither the exact formula nor an approximation for P_D is known, it will be estimated using MC simulations.

2.5.1 False Alarm Probabilities, Thresholds, and Gains

To estimate the FAP of both proposed detectors, we use g-method which is previously described in chapter 1. The primary data $\mathbf{x}$ and training data vectors $\mathbf{x}(l)$ are jointly independent and complex Gaussian, sharing a true covariance matrix $\mathbf{R}$. Because GM-STAP and E-AMF have CFAR property, we can set $\mathbf{R} = \mathbf{I}$ to simulate the performance of both detectors. In estimating the sample covariance matrix for FAP estimation, the training data vectors are assumed free of target signal.

FAP of GM-STAP detector

To evaluate the FAP of GM-STAP, for the sake of simplicity, we can write the test in (2.1) as[4]

$$|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{x}|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \eta_g^2 \left(\prod_{l=1}^{L} |\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{x}(l)|^2 \right)^{1/L} \tag{2.16}$$

[4]The LHS is, conditioned on the covariance estimate $\widehat{\mathbf{R}}$, an exponential random variable.

Using above test, the FAP can be expressed as

$$\begin{aligned}\alpha &= P\left(|\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}|^2 > \eta_g^2\left(\prod_{l=1}^{L}|\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}(l)|^2\right)^{1/L}\Big|H_0\right)\\ &= E\left\{P\left(|\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}|^2 > \eta_g^2\left(\prod_{l=1}^{L}|\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}(l)|^2\right)^{1/L}\Big|\mathbf{X}_L, H_0\right)\right\}\\ &\triangleq E\{g(\mathbf{X}_L)\} \qquad (2.17)\end{aligned}$$

By following same derivations as in section 1.3.2, the only changed variable is D (see (1.10)). We can redefine D for GM-STAP as

$$D_G = \frac{\left(\prod_{l=1}^{L}|\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{x}(l)|^2\right)^{1/L}}{\mathbf{s}^\dagger(\widehat{\mathbf{R}}^{-1})^2\mathbf{s}} \qquad (2.18)$$

The numerator in the equation above is the right-hand side of GM-STAP test replacing the right-hand side of AMF test as in (1.10). Therefore, $g(\mathbf{X}_L)$ becomes

$$g(\mathbf{x}_L) = e^{-\eta_g^2 D_G}$$

and the FAP expression in term of D_G is

$$\alpha = E\{e^{-\eta_g^2 D_G}\} \qquad (2.19)$$

The weighting function $W(\mathbf{X}_L;\theta)$, I-function and its derivatives are the same as described in the section 1.3.2. The threshold is estimated adaptively by using recursive formula described in (1.13).

Some simulations are run for case $L = 128$ and $N = 64$, by applying the g-method. The adaptive IS and inverse IS method are run simultaneously to estimate the optimal biasing parameters θ and thresholds. The scaling is used as biasing technique. We simulate FAP from 10^{-1} to 10^{-7} with logarithmic steps of one decade.

We can see the threshold estimation by using adaptive inverse IS method in the Figure 2.1. For each FAP, the threshold reaches its convergence to optimal value in 20 iterations. By using these thresholds we get FAP estimates pictured in Figure 2.2. The convergence of biasing parameter θ of each FAP to its optimal value can be seen in the Figure 2.3. From this figure we can observe that the biasing parameter for each FAP estimate is very close to unity. This means that the metric D_G has small variance, probably caused by the choice of L and N.

Thresholds (GM)

Figure 2.4 shows the thresholds η_g for corresponding FAP. This can be approximated by a linear interpolation between η_g and $\log_{10}\alpha$ which can be expressed mathematically by the relation

$$\eta_g = -1.3\log_{10}\alpha + 3.2$$

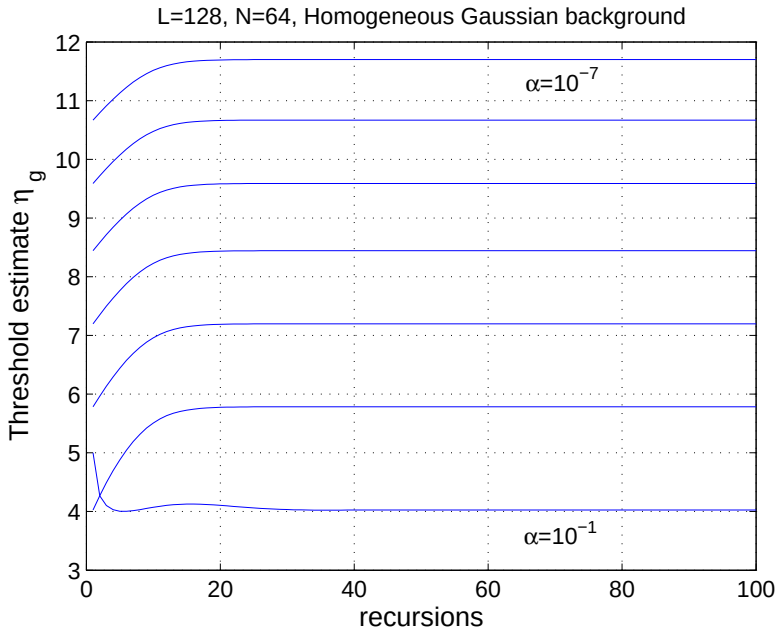

Figure 2.1: Convergence of η_g. Inverse IS method for GM-STAP detector.

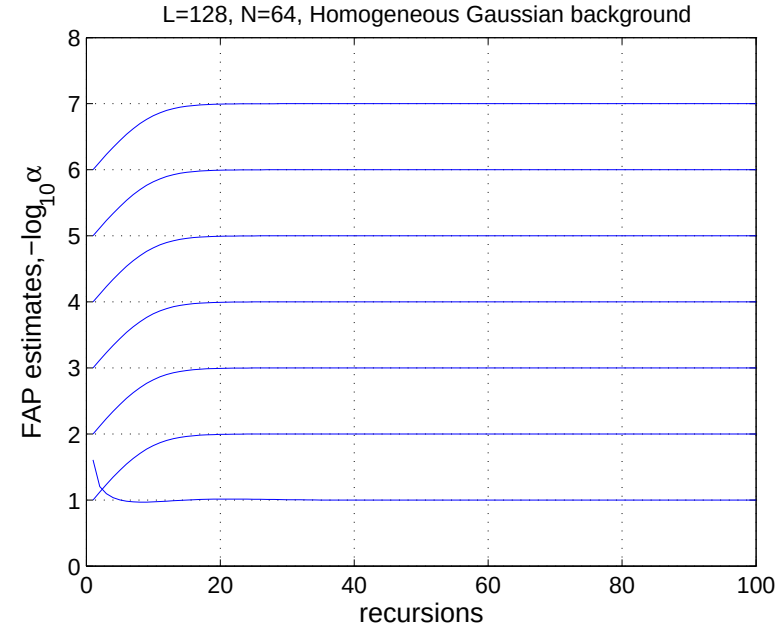

Figure 2.2: FAP using IS for GM-STAP detector.

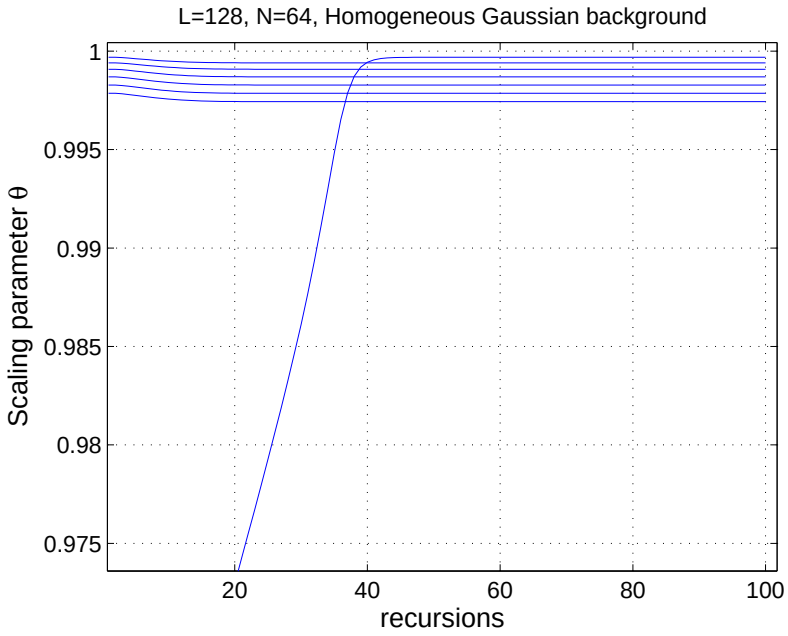

Figure 2.3: Biasing parameter θ for GM-STAP detector.

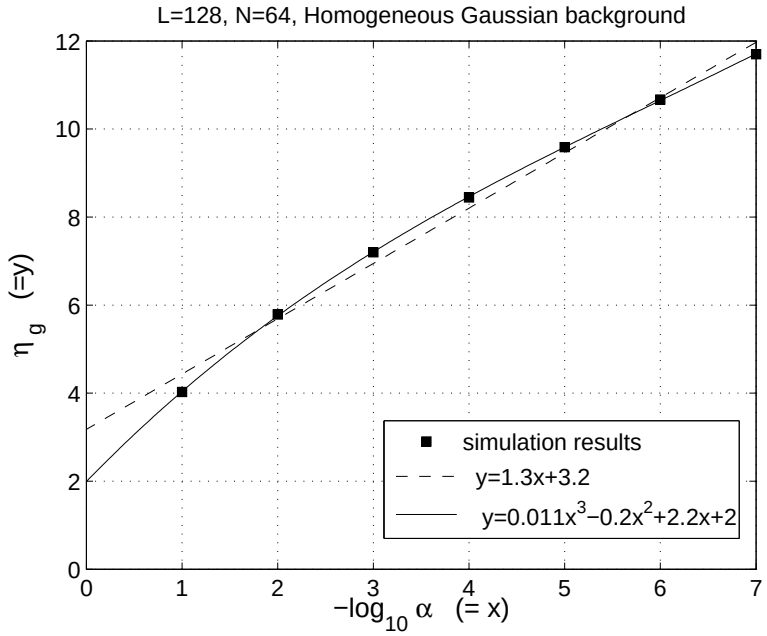

Figure 2.4: η vs. FAP for GM-STAP.

The gradient of the linear graph is 1.3 which means if we decrease the FAP by geometric factor 10^{-1} then the threshold will increase by arithmetic factor 1.3. A more accurate relation between η_g and $\log_{10}\alpha$ can be represented by the cubic interpolation

$$\boxed{\eta_g = -0.011\,(\log_{10}\alpha)^3 - 0.2\,(\log_{10}\alpha)^2 - 2.2\,\log_{10}\alpha + 2} \tag{2.20}$$

To be noted is that this interpolation has been obtained for the FAP range $10^{-1} - 10^{-7}$.

Simulation gains

The simulation gains for different FAPs can be seen in Figure 2.5. The obtained gains are quite high but lower compared to AMF ($L = 128, N = 64$) and much lower compared to AMF for $L = 704$ and $N = 352$. We can expect that for higher L and N, the simulation gains of GM-STAP will be higher. However, using IS in simulating the FAP of GM-STAP indeed reduces the computational load. As an example, for FAP 10^{-6}, the trials needed are only about 2300 while MC needs 10^8.

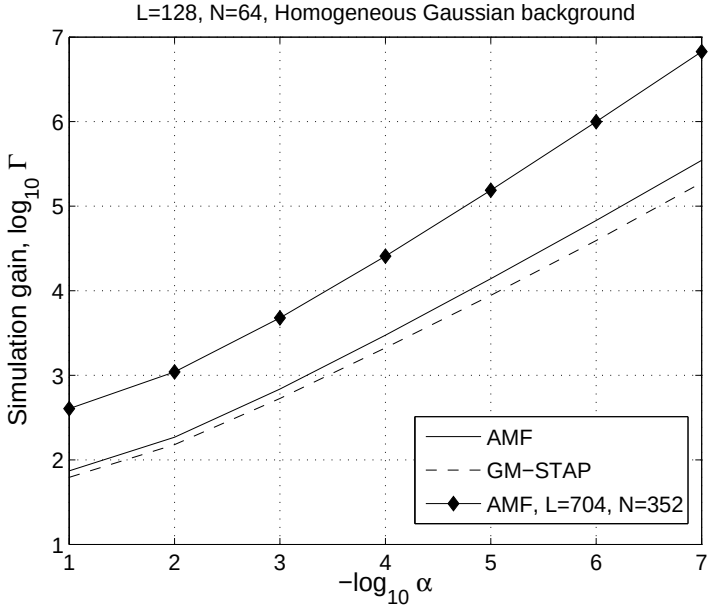

Figure 2.5: Simulation gains Γ vs FAP for three detectors.

FAP of E-AMF

The false alarm probability of E-AMF is also estimated by using IS and g-method. To derive the g-method estimator, for the sake of simplicity, we can

express (2.2) as

$$\begin{aligned} |\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}|^2 &\underset{H_0}{\overset{H_1}{\gtrless}} \eta_e^2 \left(\frac{1}{L} \sum\nolimits_{l=1}^{L} |\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}(l)| \right)^2 \\ |\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}|^2 &\underset{H_0}{\overset{H_1}{\gtrless}} \eta_e^2 Z \end{aligned} \tag{2.21}$$

where

$$Z \equiv \left(\frac{1}{L} \sum\nolimits_{l=1}^{L} |\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}(l)| \right)^2$$

The FAP for E-AMF then can be written as

$$\begin{aligned} \alpha &= P(|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{X}|^2 > \eta_e^2 Z | H_0) \\ &= E\{P(|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{X}|^2 > \eta_e^2 Z | \mathbf{X}_L, H_0)\} \\ &\triangleq E\{g(\mathbf{X}_L)\} \end{aligned} \tag{2.22}$$

By carrying out the same steps as in section 1.3.2, we get the expression for $g(\mathbf{X}_L)$ as

$$g(\mathbf{X}_L) = e^{-\eta_e^2 D_E}$$

where

$$D_E = \frac{Z}{\mathbf{s}^\dagger (\widehat{\mathbf{R}}^{-1})^2 \mathbf{s}}$$

The expression of FAP becomes

$$\begin{aligned} \alpha &= E\{e^{-\eta_e^2 D_E}\} \\ &= E_\star\{e^{-\eta_e^2 D_E} W(\mathbf{X}_L;\theta)\} \end{aligned}$$

and its estimate is

$$\widehat{\alpha}_g = \frac{1}{K} \sum\nolimits_{1}^{K} e^{-\eta_e^2 D_E} W(\mathbf{X}_L;\theta); \quad \sim f_\star$$

The variables needed for adaptive IS like scaling parameter θ,weighting function $W(\mathbf{X}_L;\theta)$, I-function and its derivatives are the same as described in section 1.3.2. To find the optimal biasing parameter θ we used Newton recursive algorithm described in (1.12). The thresholds for each FAP are adaptively estimated by using inverse IS. We may reformulate the adaptive inverse IS algorithm described in (1.13) and (1.14) for the E-AMF case as

$$\eta_{m+1} = \eta_m + \delta_\eta \frac{\alpha_o - \widehat{\alpha}(\eta_m)}{\widehat{\alpha'}(\eta_m)}, \quad m = 1, 2, \ldots \tag{2.23}$$

where δ_η is a step-size parameter and the derivative estimator in the denominator is given by

$$\widehat{\alpha'}(\eta_m) = -\frac{2\,\eta_m\, D_E}{K} \sum\nolimits_{1}^{K} e^{-\eta_m D_E} W(\mathbf{X}_L;\theta); \quad \sim f_\star$$

$$= \quad -2\,\eta_m\,D_E\,\widehat{\alpha}(\eta_m) \tag{2.24}$$

with the prime indicating derivative with respect to η_m. The symbols η_m and m correspond to the η_e and iteration number respectively.

We test adaptive IS and inverse IS algorithms simultaneously for case $L = 128$ and $N = 64$ to estimate FAPs from 10^{-1} to 10^{-7}. The estimated FAPs are described in Figure 2.7 by using thresholds found from optimization process pictured in Figure 2.6. We can see that the adaptive inverse IS algorithm converges in 20 iterations which is quite fast. The biasing parameter (θ) optimization process also converges in 20 iterations as pictured in Figure 2.8. As in case of GM-STAP detector, the biasing parameters for E-AMF are also close to unity and has small spread.

Thresholds (E-AMF)

The relation between the thresholds and -$\log_{10}\alpha$ is considerably more linear in this case with gradient 1, as described in Figure 2.9 and expressed as

$$\eta_e = -\log_{10}\alpha + 2.7$$

A more accurate relation is expressed in term of the cubic equation

$$\boxed{\eta_e = -0.01\,(\log_{10}\alpha)^3 - 0.17\,(\log_{10}\alpha)^2 - 1.9\,(\log_{10}\alpha) + 1.7} \tag{2.25}$$

By using this expression we may have rough estimates of the thresholds for lower FAPs.

Simulation gains

Figure 2.10 shows the simulation gains by implementing IS simulation with g-method. We can observe that simulation gains of E-AMF are slightly higher than for GM at high FAPs and almost the same at low FAPs. Compared to the AMF, the gains of E-AMF are almost the same at high FAPs but lower at low FAPs. As in case of AMF, larger values of L and N may result higher in simulation gains.

2.5.2 P_D in homogeneous case

In this subsection we present the detection probability of GM-STAP and E-AMF detectors in homogeneous background and compare them with the AMF. We set FAP to 10^{-6}, the value of L and N are the same as in FAP simulations. All the simulations concerning detection probability are done by using standard Monte Carlo (MC) technique, because the probabilities are not very low. The number of samples are kept larger than $100/P_D$. The detection probability in case of non-fluctuating target can be seen in the Figure 2.11. It shows that the detection probability of AMF is better than GM-STAP and E-AMF in the homogeneous case. For P_D equals to 0.5, E-AMF has about 0.18 dB detection loss compare to AMF, while GM-STAP suffers 0.3 dB loss.

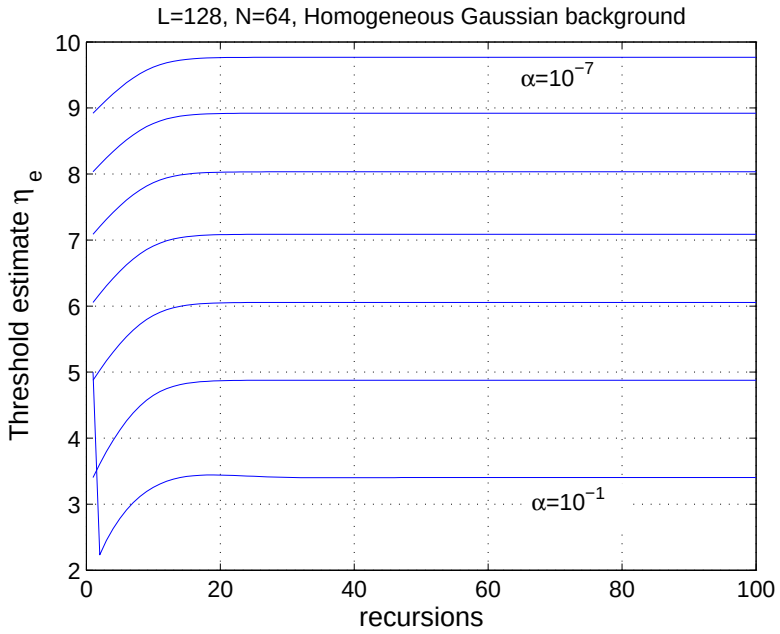

Figure 2.6: Convergence of η_g. Inverse IS method for E-AMF.

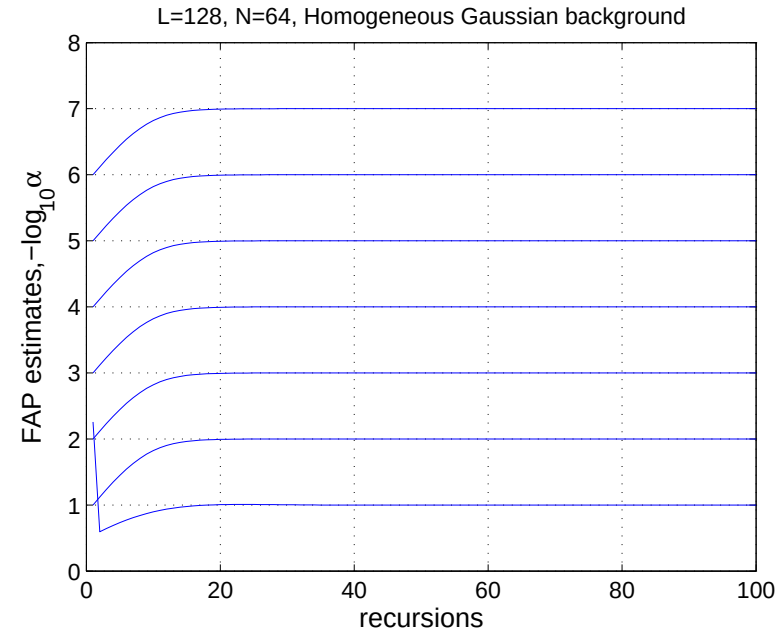

Figure 2.7: Convergence of FAP using IS for E-AMF.

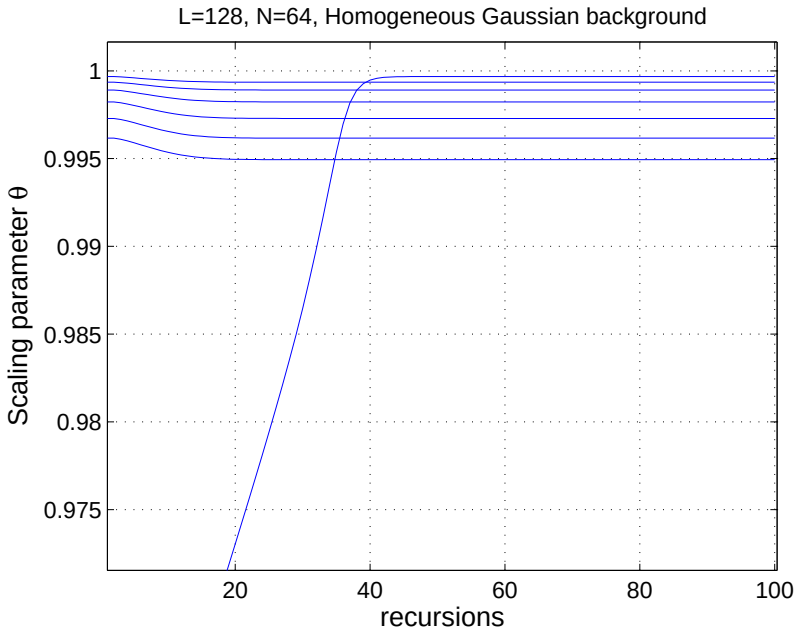

Figure 2.8: Adaptive IS biasing θ for E-AMF.

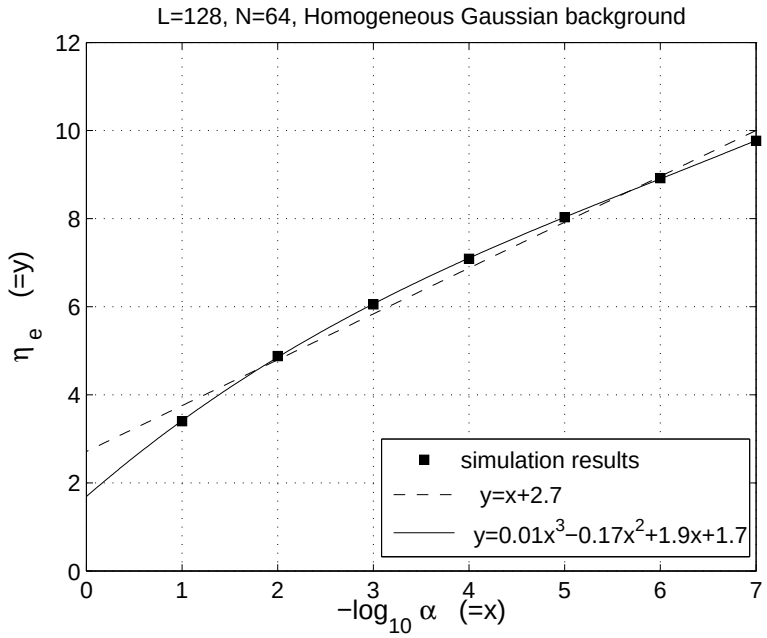

Figure 2.9: η vs. FAP for E-AMF.

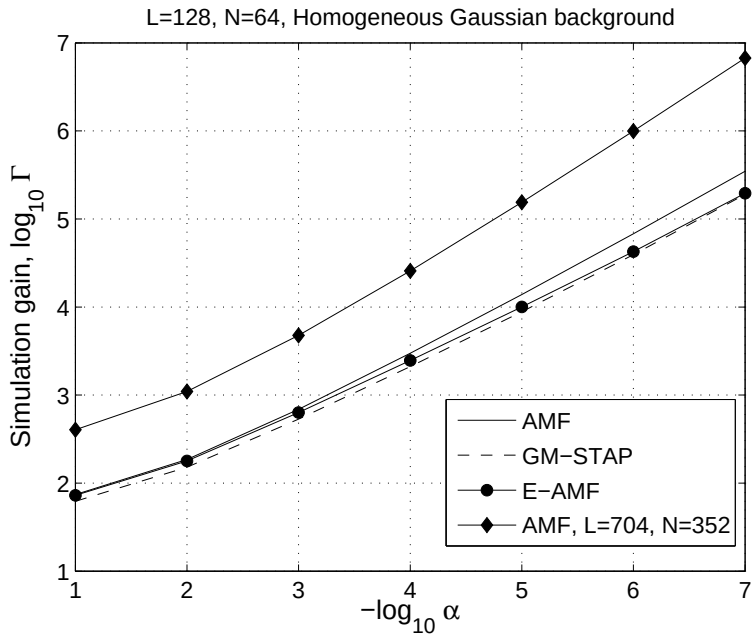

Figure 2.10: Simulation gains Γ vs FAP for three detectors.

We also do some simulations to estimate detection probability when the Swerling I fluctuating target model is included. To simulate the performance because of this fluctuating target model, a combination of g-method and MC is used. This is done to obtain smoother results. The result is presented in Figure 2.12. We can observe that for P_D equals to 0.5, GM-STAP has loss about 0.3 dB and E-AMF has loss about 0.15 dB compared to AMF.

From these simulation results we can conclude that in homogeneous case, regardless the fluctuation of the target, both proposed detectors have small detection loss compare to AMF.

2.5.3 P_D in the presence of nonhomogenities

In this subsection we will show the simulation results that compare the performances of GM-STAP and E-AMF with AMF in the presence nonhomogenities. The nonhomegenity we consider here is the presence of interfering targets. The power of the interfering targets are taken to be the same as the power of the actual target[5]. Figure 2.13 shows the comparison of the performance of GM-STAP and AMF in the presence of 2 interferers and the target signal is constant,i.e. non-fluctuating. At $P_D = 0.5$, AMF has detection loss about 0.98 dB w.r.t

[5]It is assumed that interferers have the same angle-Doppler properties as those of the primary target.

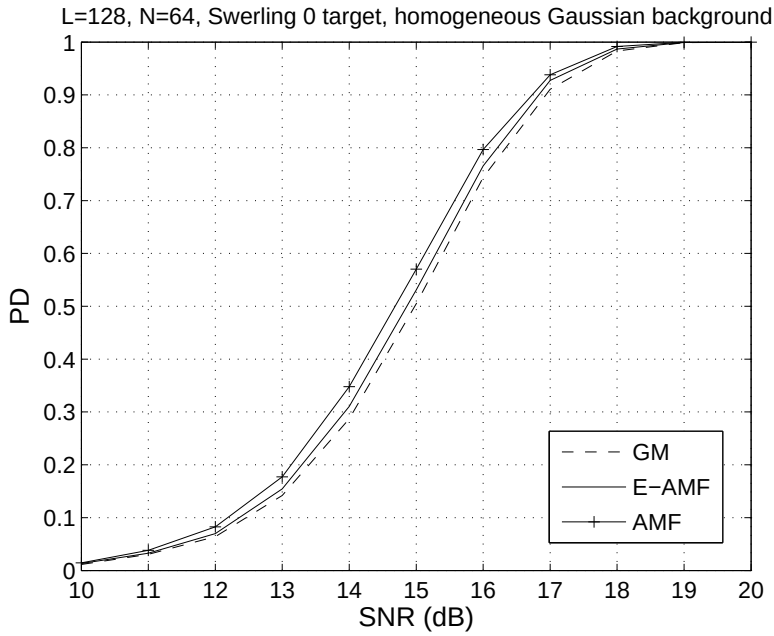

Figure 2.11: P_D in homogeneous case, Swerling 0 target.

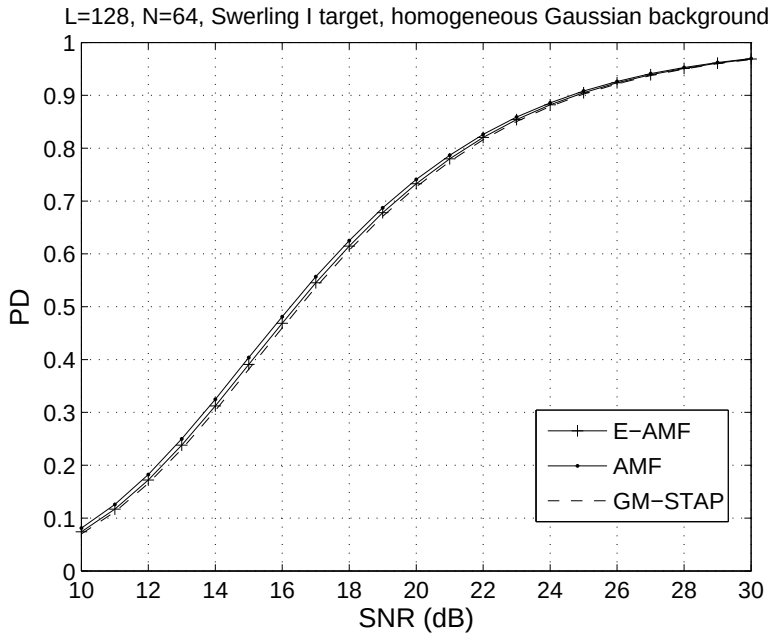

Figure 2.12: P_D in homogeneous case, Swerling I target.

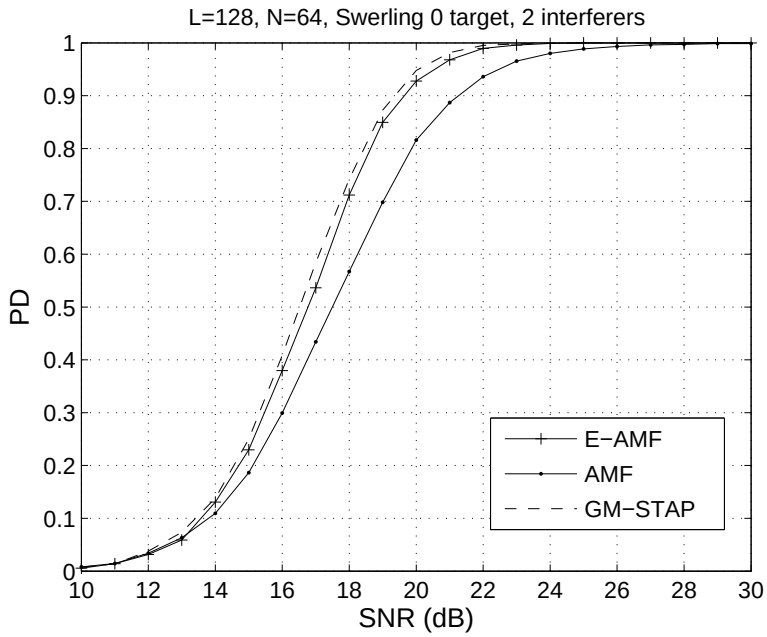

Figure 2.13: P_D of AMF, GM-STAP, and E-AMF.

GM-STAP and about 0.73 dB compared to E-AMF. This detection loss will increase for higher P_D. For example at $P_D = 0.9$, AMF suffers loss about 1.9 dB w.r.t GM and 1.6 dB w.r.t E-AMF. From the graph, we can also see that the detection loss of E-AMF compared to GM is quite constant i.e about 0.25-0.3 dB.

For case 3 interferers, the losses of AMF are even higher. At $P_D = 0.5$, AMF has loss about 2.5 dB w.r.t GM and 1.9 dB w.r.t to E-AMF. For an extreme case, i.e. at $P_D = 0.9$, AMF has loss about 9 dB and 8 dB compared to GM and E-AMF respectively. The E-AMF itself has loss 0.58 dB (at $P_D = 0.5$) and 0.85 (at $P_D = 0.9$) dB compared to GM.

These facts show the robustness of GM and E-AMF detectors compared to AMF. In the case of Swerling I fluctuating target model, both detectors still show their robustness compared to AMF as can be seen in the Figure 2.15. We can conclude that in the presence of interfering targets, GM-STAP performs the best while AMF is the worst. The E-AMF detector also shows its robutness in these simulations and has small detection loss compared to GM detector.

2.6 Conclusion

In this chapter we introduced the GM-STAP and E-AMF detectors. We show that IS methods work well in estimating the FAP and determining the thresholds

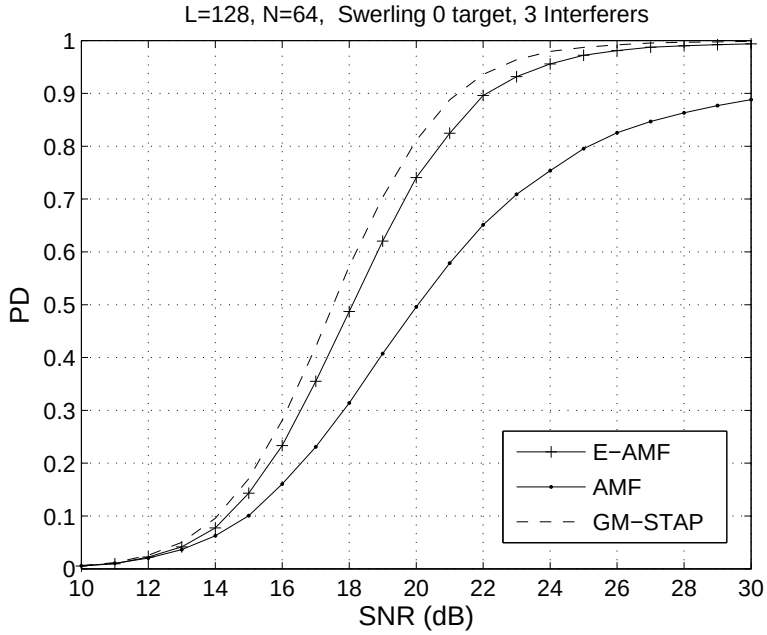

Figure 2.14: P_D of AMF, GM-STAP, and E-AMF.

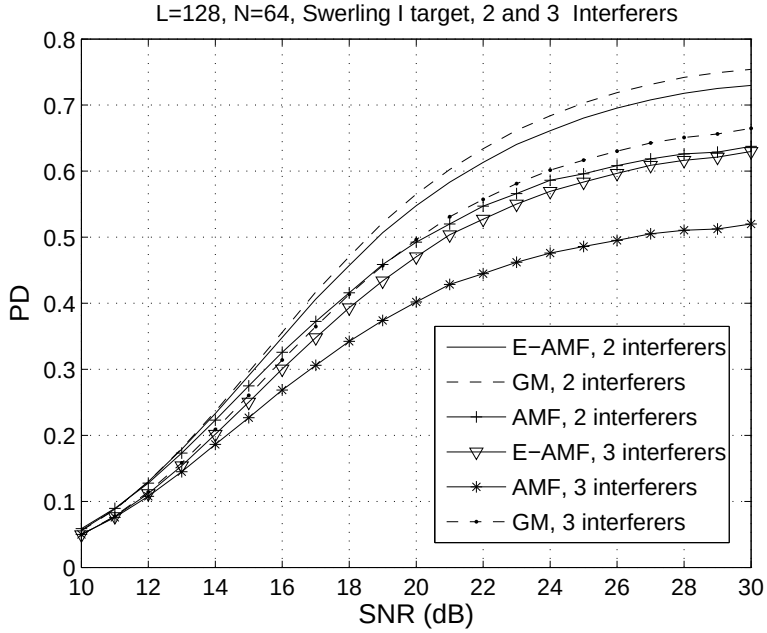

Figure 2.15: P_D of AMF, GM-STAP, and E-AMF.

Table 2.1: Comparison of three STAP detectors.

Q	AMF-S	AMF-E	GM-STAP
Is it CFAR?	Yes	Yes	Yes
homogeneous P_d	-	very small loss	very small loss
Robust to interferers?	No	somewhat	most

of GM-STAP and E-AMF detectors. However, IS simulation gains for GM-STAP and E-AMF are not very high, and they can be increased by developing better biasing methods which is subject for future research.

Simulation results show that the GM-STAP and E-AMF detectors have quite small loss in detection probability compared to AMF, but they have very good performance and show robustness in the presence of interfering targets. The AMF fails to maintain its robustness in this situation. These findings are summarized in Table 2.1.

Chapter 3

Further results on IS for STAP detection

This third report (or chapter) begins with some additional results on the E-AMF and GM-STAP detectors for the case of $L = 704$ and $N = 352$. These complete our investigations into the two detectors and complement parallel results that were obtained previously for the square-law AMF detector with the above L and N values. Comparisons with the latter detector have also been made.

The results are contained in Figures 3.1 - 3.5 for the E-AMF detector. Inverse IS results are in Fig. 3.1 and 3.2. Estimated optimum biasing parameters and the I-functions are in Fig. 3.3 and 3.4 respectively. IS gains estimated during simulation are shown in Fig. 3.5. Corresponding results for the GM-STAP detector are in Figures 3.6 - 3.8. Thresholds for the 2 detectors (including AMF) as functions of FAPs are in Fig 3.9. The cubic interpolations

$$\eta_E = 0.01x^3 - 0.18647x^2 + 1.8822x + 1.725467$$

and

$$\eta_G = 0.011813x^3 - 0.21985x^2 + 2.22917x + 2.040834$$

where $x = -\log_{10}\alpha$ can be used to determine thresholds for the E-AMF and GM detectors in the FAP range 10^{-1} - 10^{-7} respectively. Estimated IS gains as functions of FAPs are shown in Fig 3.10. It is noted that simulation gains for the 3 detectors are of the same order.

3.1 General remarks on IS

Point of application of biasing

Some comments of a general nature regarding application of IS to signal processing algorithms are made here. They give some insight into the thinking behind our efforts to simulate the above mentioned algorithms. The simulation procedures described thus far in the previous chapters are applicable to any detector.

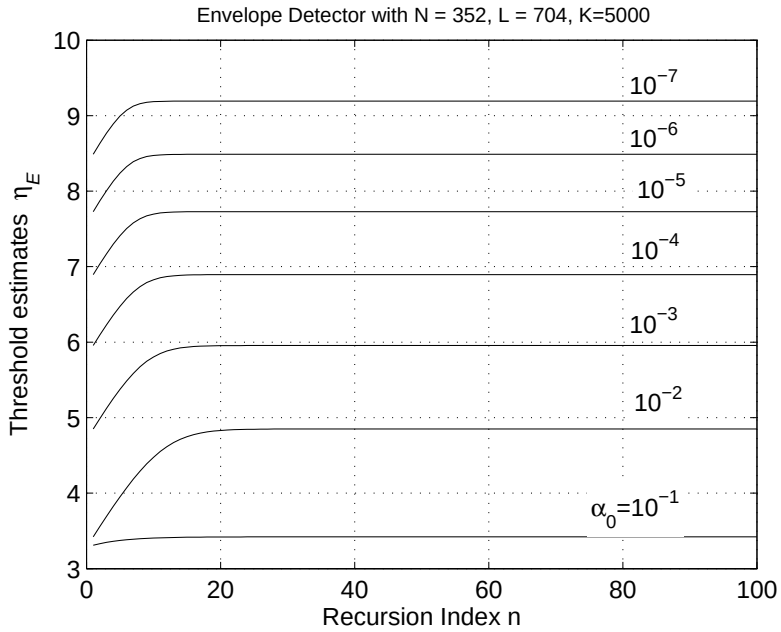

Figure 3.1: Inverse IS thresholds.

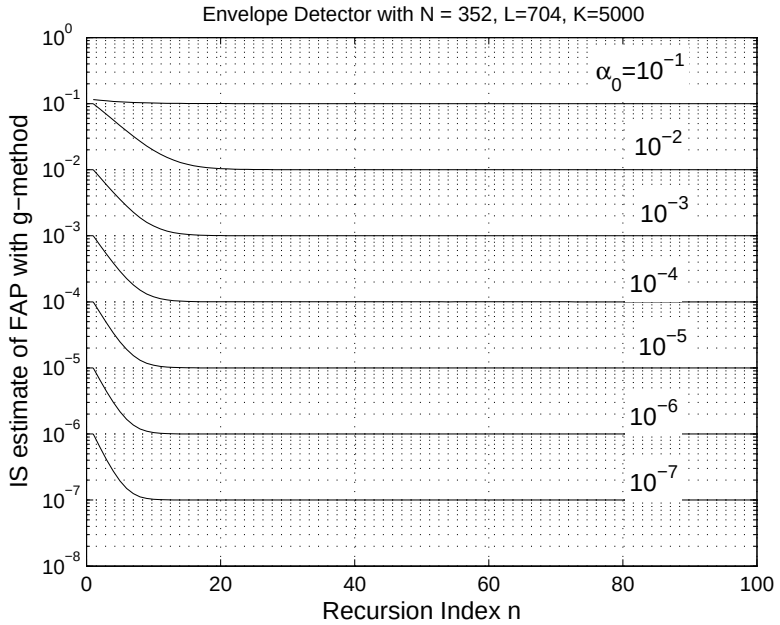

Figure 3.2: FAP estimates.

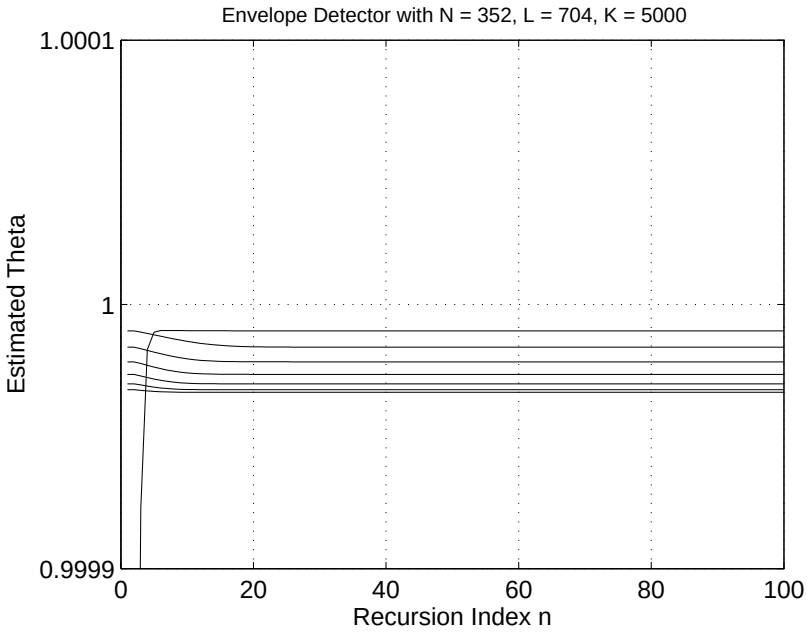

Figure 3.3: Optimum scaling parameters.

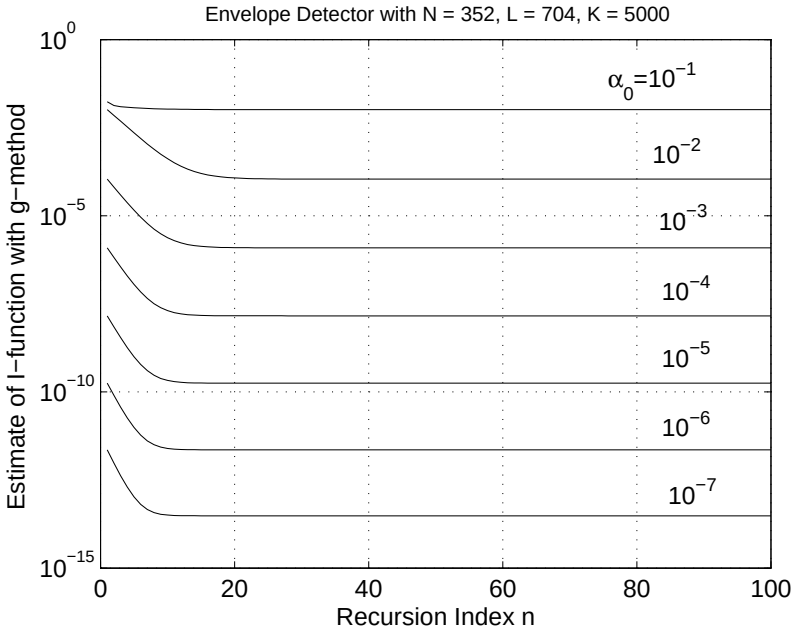

Figure 3.4: I-function estimates.

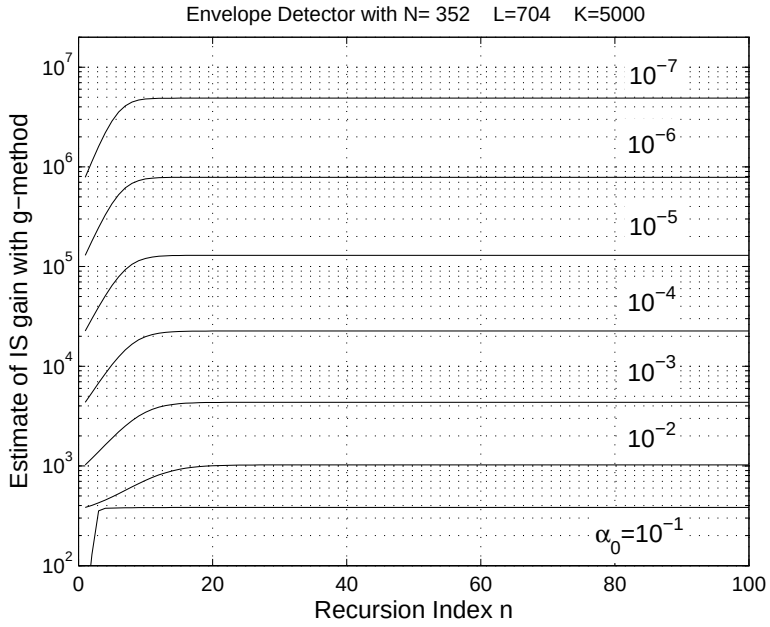

Figure 3.5: IS gain estimates.

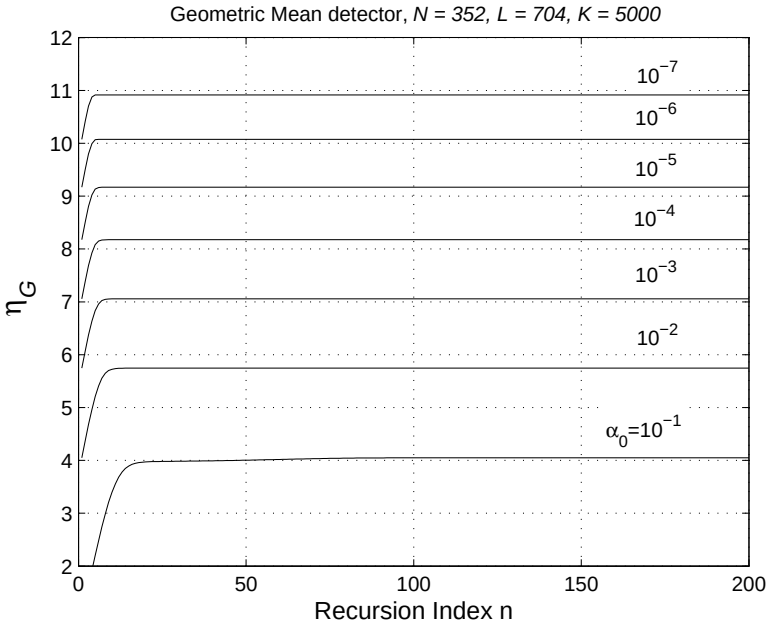

Figure 3.6: Inverse IS thresholds.

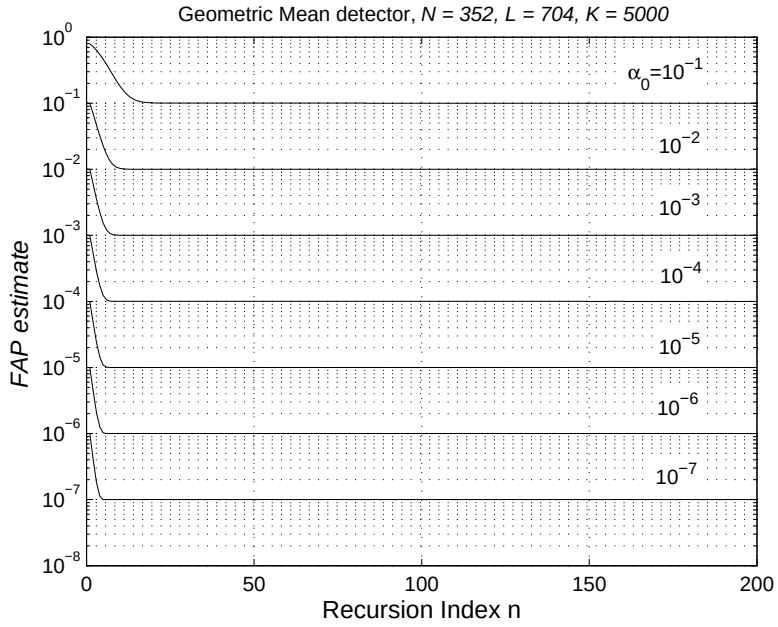

Figure 3.7: FAP estimates.

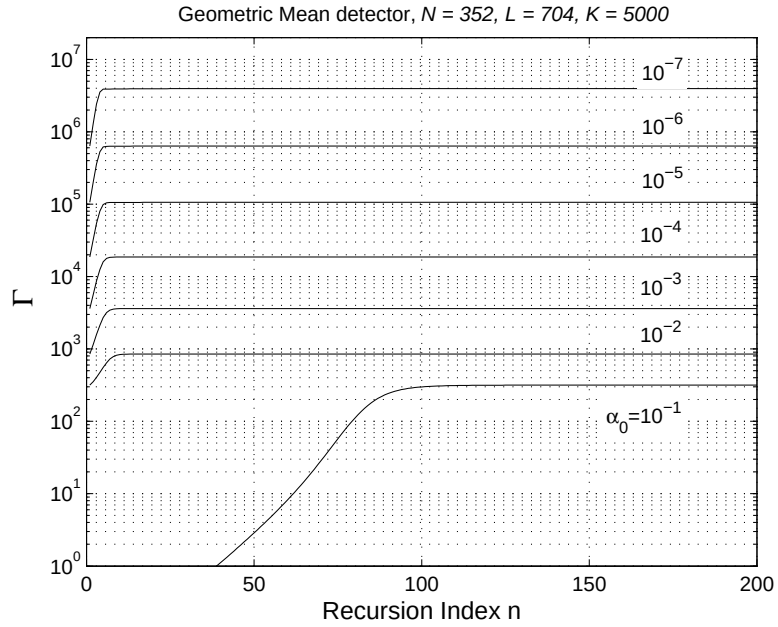

Figure 3.8: IS gain estimates.

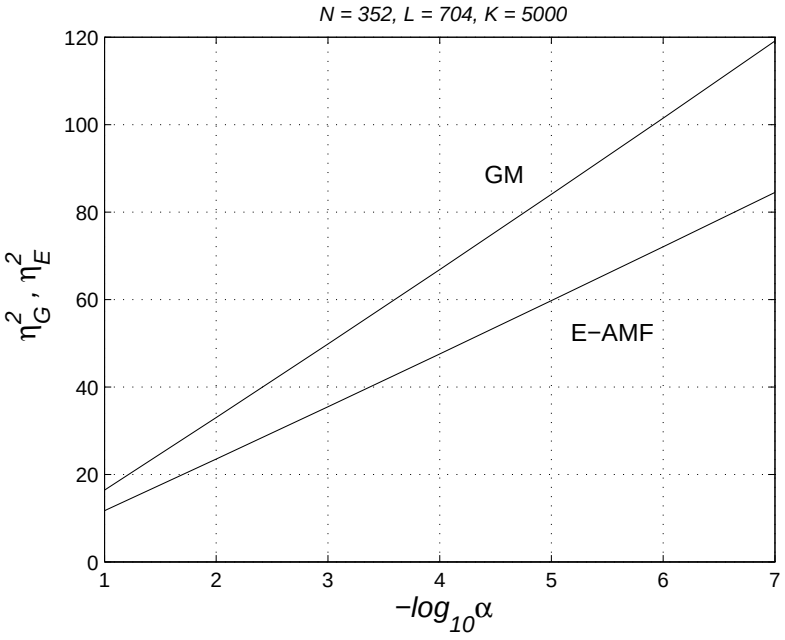

Figure 3.9: Thresholds as functions of FAP.

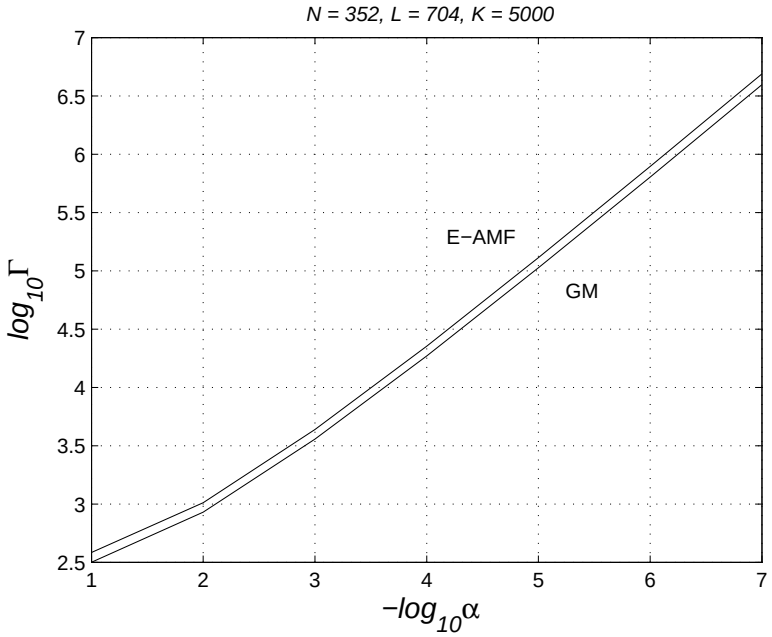

Figure 3.10: Gain comparisons as functions of FAP.

The IS biasing is performed on the input random variables and we refer to this as *input biasing*. Accuracy of the IS estimates and resulting simulation gains will of course depend on the particular detection algorithm under study. It is generally true (and intuitive) that better estimator performance can be obtained if IS biasing can be carried out closer to the point in the processing chain of the detector where the actual (rare event) decisionmaking is done. This of course necessitates knowledge of density functions of the processes at the point where biasing is to be implemented. Often, input stochastic variables may have undergone transformations whose results are difficult to characterize statistically in analytical form, and we have to rely on the general method above. However, when such transformations can be characterized, then IS should be carried out using the modified processes. Therefore, it is desirable to perform biasing and IS as close to the final decisionmaking point as permitted by availability of knowledge of probability density functions.

Another approach to biasing that is sometimes possible is to perform a series of (linear and/or nonlinear) transformations of the input processes as if one were carrying out a mathematical analysis of the algorithm. The transformations are carried out until the point beyond which it may not be possible to determine the density functions of the transformed processes without considerable mathematical effort. Biasing is then performed at this stage, the hope being that the g-method becomes applicable. This procedure may produce higher simulation gains than simple input biasing which, of course, is the easiest to implement. It is illustrated in some of the following examples.

3.2 The NMF STAP detector

The use of IS to characterize the normalized matched filter (NMF) detector and some of its variants is discussed in the rest of this chapter. Expressions for the FAP and detection probability of the NMF detector operating in homogeneous Gaussian clutter are known in closed form. However, much can be learnt about biasing for this class of detectors from the exercise of applying IS to estimate its FAP. The variants are the normalized adaptive matched filter (NAMF) and the low-rank NAMF detectors. All these detection algorithms have been treated in detail in [3], [15], and other papers.

The NMF detection test is given by

$$\Lambda_{\text{NMF}} \equiv \frac{|\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{x}|^2}{(\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s})(\mathbf{x}^\dagger \mathbf{R}^{-1}\mathbf{x})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{3.1}$$

following the usual notations. Its FAP in Gaussian interference is known, [15], and particularly easy to derive. It differs from the matched filter (for known interference covariance matrix $\mathbf{R}$) in the normalization term which is the second one in the denominator. The FAP of the detector is given by

$$\alpha_{\text{NMF}} = (1-\eta)^{N-1}$$

and it has the CFAR property of being invariant to the interference covariance matrix $\mathbf{R}$. Furthermore, as is evident from (3.1), the FAP is also invariant to any scaling of the primary and secondary data.

3.2.1 The FAP of the NMF detector (*using IS*)

We begin with a simple (re-) derivation of the FAP of the detector. It helps to illustrate an elegant aspect of IS, by which it is sometimes possible to derive a perfect estimate of a rare-event probability. Furthermore, although this problem can be analyzed completely, it is an application of the procedure described in the last paragraph of the introduction.

It is assumed that $\mathbf{x} \sim \mathcal{CN}_N(0, \mathbf{R})$. With a whitened data vector defined as $\mathbf{x}_1 \equiv \mathbf{R}^{-1/2}\mathbf{x}$ and a transformed steering vector $\mathbf{s}_1 \equiv \mathbf{R}^{-1/2}\mathbf{s}$, the test statistic of (3.1) takes the form

$$\begin{aligned}
\Lambda_{\mathrm{NMF}} &= \frac{|\mathbf{s}_1^\dagger \mathbf{x}_1|^2}{(\mathbf{s}_1^\dagger \mathbf{s}_1)(\mathbf{x}_1^\dagger \mathbf{x}_1)} \\
&= \frac{\mathbf{x}_1^\dagger \mathbf{s}_1}{\|\mathbf{s}_1\|} \frac{\mathbf{s}_1^\dagger \mathbf{x}_1}{\|\mathbf{s}_1\|} \frac{1}{\mathbf{x}_1^\dagger \mathbf{x}_1} \\
&= \frac{\mathbf{x}_1^\dagger \mathbf{u}_1 \mathbf{u}_1^\dagger \mathbf{x}_1}{\mathbf{x}_1^\dagger \mathbf{x}_1} \\
&= \frac{y_1^\star y_1}{\mathbf{x}_1^\dagger \mathbf{x}_1} \\
&= \frac{|y_1|^2}{\|\mathbf{x}_1\|^2}
\end{aligned} \tag{3.2}$$

where $\mathbf{u}_1 = \mathbf{s}_1/\|\mathbf{s}_1\|$ is an N-dimensional unit vector and $y_1 = \mathbf{u}_1^\dagger \mathbf{x}_1$ a scalar random variable. Using for example a Gram-Schmidt procedure, an N-dimensional basis can be formed by determining $N-1$ other unit vectors in the orthogonal subspace of $\mathbf{u}_1$. Denoting the former by $\mathbf{u}_i$, $i = 2, \ldots, N$, we define the corresponding random variables $y_i = \mathbf{u}_i^\dagger \mathbf{x}_1$. Then

$$\begin{aligned}
|y_1|^2 + \textstyle\sum_{i=2}^N |y_i|^2 &= \sum_{i=1}^N |y_i|^2 \\
&= \mathbf{x}_1^\dagger \sum_{i=1}^N \mathbf{u}_i \mathbf{u}_i^\dagger \, \mathbf{x}_1 \\
&= \mathbf{x}_1^\dagger \mathbf{I} \, \mathbf{x}_1 \\
&= \|\mathbf{x}_1\|^2
\end{aligned} \tag{3.3}$$

with the $\{y_i\}_1^N$ being i.i.d. and distributed as $\mathcal{CN}_1(0,1)$. The test therefore becomes

$$\Lambda_{\rm NMF} = \frac{|y_1|^2}{|y_1|^2 + \sum_{i=2}^N |y_i|^2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{3.4}$$

which can be put in the form

$$\frac{u}{\sum_{i=2}^N v_i} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_o \equiv \frac{\eta}{1-\eta} \tag{3.5}$$

where

$$u \equiv |y_1|^2 \quad \text{and} \quad v_i \equiv |y_i|^2, \quad i = 2, \ldots, N$$

The FAP of the NMF detector is then

$$\begin{aligned} \alpha_{\rm NMF} &= P\big(U \geqslant \eta_o \sum\nolimits_{i=2}^N V_i\big) \\ &= E\Big\{P\big(U \geqslant \eta_o \sum\nolimits_{i=2}^N V_i \Big| V_2, \ldots, V_N\big)\Big\} \end{aligned} \tag{3.6}$$

The first line above has exactly the same form as the FAP of a CA-CFAR detector as U and each V_i are i.i.d. (unit) exponential random variables; the formula for this probability is well known. To re-derive the latter, a g-method estimator can be written as

$$\widehat{\alpha}_{\rm NMF} = \frac{1}{K}\sum_1^K \mathrm{e}^{-\eta_o \sum_2^N V_i} \cdot W(V_2, \ldots, V_N) \tag{3.7}$$

Scaling each V_i with a and using the resulting weighting function

$$W(v_2, \ldots, v_N) = a^{N-1}\, \mathrm{e}^{-(1-1/a)\sum_2^N v_i} \tag{3.8}$$

yields the estimator

$$\widehat{\alpha}_{\rm NMF} = \frac{1}{K}\sum_1^K a^{N-1}\, \mathrm{e}^{-(\eta_o+1-1/a)\sum_2^N V_i} \tag{3.9}$$

If the scaling factor is chosen as $a = 1/(1+\eta_o)$ in this estimator, then

$$\begin{aligned} \widehat{\alpha}_{\rm NMF} &= \frac{1}{K}\sum_1^K \left(\frac{1}{1+\eta_o}\right)^{N-1} \\ &= (1-\eta)^{N-1}, \text{ a constant} \\ &= \alpha_{\rm NMF} \end{aligned} \tag{3.10}$$

the last step following from the fact that the variance of this unbiased estimator is zero.

3.2.2 FAP estimation by IS: rotation of primary vector

If we want to estimate $\alpha_{\rm NMF}$ using input biasing, then it is clear from the test statistic in (3.1) that a simple scaling of the elements of the primary data vector $\mathbf{x}$ will be useless. A form of biasing can however be developed using the (well known) fact that the test statistic is actually a cosine-squared one, the concerned angle being that between the transformed steering (or unit) and whitened primary data vectors $\mathbf{s}_1$ (or $\mathbf{u_1}$) and $\mathbf{x}_1$ respectively, as evident from (3.2). The frequency of false alarm events in a simulation can therefore be increased by biasing $\mathbf{x}_1$ so as to decrease the angle toward zero or increase it toward π. This can be accomplished by a rotation of $\mathbf{x}_1$ which, since it is being assumed for this detector that the actual data covariance matrix $\mathbf{R}$ is completely known, is equivalent to input biasing (of the primary data vector $\mathbf{x}$).

Consider biasing the whitened primary data vector $\mathbf{x}_1$ by rotation with an $N \times N$ matrix $\mathbf{A}$. The biased vector is $\mathbf{A}\mathbf{x}_1$ and will be distributed as $\mathcal{CN}_N(0, \mathbf{R}_\star)$ with covariance matrix $\mathbf{R}_\star$ given by $\mathbf{R}_\star = \mathbf{A}\mathbf{A}^\dagger$, since $\mathbf{x}_1 \sim \mathcal{CN}_N(0, \mathbf{I})$. The weighting function is then given by

$$\begin{aligned} W(\mathbf{x}_1; \mathbf{A}) &= \frac{f(\mathbf{x}_1)}{f_\star(\mathbf{x}_1)} \\ &= |\mathbf{R}_\star| \, \exp\big(-\mathbf{x}_1^\dagger (\mathbf{I} - \mathbf{R}_\star^{-1}) \mathbf{x}_1 \big) \end{aligned} \tag{3.11}$$

where $|\cdot|$ denotes matrix determinant. The FAP estimator for the test in (3.2) then takes the form

$$\widehat{\alpha}_{\rm NMF} = \frac{1}{K} \sum_1^K 1\big(|\mathbf{u}_1^\dagger \mathbf{x}_1|^2 \geq \eta \, \|\mathbf{x}_1\|^2 \big) \, W(\mathbf{x}_1; \mathbf{A}); \quad \sim f_\star \tag{3.12}$$

The problem now centers around determining an effective rotation or biasing matrix $\mathbf{A}$. There is clearly an unbounded number of rotation matrices to choose from. To narrow the choice, it is most convenient to make the matrix dependent on a single (biasing) parameter, say a, and search for the optimum value of the latter by minimizing the associated I-function which is given by

$$\begin{aligned} I(a) &= E_\star\big\{ 1\big(\mathcal{A}\big) \, W^2(\mathbf{x}_1; \mathbf{A}) \big\} \\ &= E\big\{ 1\big(\mathcal{A}\big) \, W(\mathbf{x}_1; \mathbf{A}) \big\} \end{aligned} \tag{3.13}$$

where $\mathcal{A} \equiv 1\big(|\mathbf{u}_1^\dagger \mathbf{x}_1|^2 \geq \eta \, \|\mathbf{x}_1\|^2 \big)$ is the false alarm event. In actual simulation we will determine a minimizer for an estimate of $I(a)$, that is

$$a_{\rm opt} = \arg\min_a \Big(\widehat{I}(a) = \frac{1}{K} \sum_1^K 1\big(\mathcal{A}\big) \, W^2(\mathbf{x}_1; \mathbf{A}); \sim f_\star \Big) \tag{3.14}$$

The form or structure of $\mathbf{A}$ has to be now decided. We assume that $\mathbf{A} = \mathbf{I} + a\mathbf{T}$ for $0 \leq a < 1$ where $\mathbf{T}$ is some $N \times N$ matrix. The biasing matrix $\mathbf{A}$ is to be

constructed so as to introduce correlation amongst the components of $\mathbf{x}_1$ in a controlled manner. To understand how this can be achieved, consider first the simple case where the unit vector $\mathbf{u}_1$ is given by

$$\mathbf{u}_1 = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ . \\ . \\ 1 \end{pmatrix}$$

This can happen if $\mathbf{R} = \mathbf{I}$ and the original steering vector $\mathbf{s}$ has all equal elements. Suppose (based on heuristic reasons) the matrix $\mathbf{T}$ is set to

$$\mathbf{T} = \begin{pmatrix} 0 & 1 & . & . & 1 \\ 1 & 0 & . & . & 1 \\ . & & & & . \\ . & & & & . \\ 1 & . & . & . & 0 \end{pmatrix}$$

Then

$$\mathbf{A} = \begin{pmatrix} 1 & a & . & . & a \\ a & 1 & . & . & a \\ . & & & & . \\ . & & & & . \\ a & . & . & . & 1 \end{pmatrix} \tag{3.15}$$

It is noted that the above matrix does not provide a *norm-preserving* rotation. We avoid simulating with $a = 1$ since this obviously leads to a singular matrix. When $a = 1$, the test statistic becomes unity and all the elements of the biased vector $\mathbf{A}\mathbf{x}_1$ are equal to $\sum_{i=1}^{N} x_{1i}$ where x_{1i}'s denote the elements of the unbiased $\mathbf{x}_1$. That is, the fully biased vector would be aligned in the direction of $\mathbf{u}_1$ or opposite to it, in $\mathcal{C}^N$. When $a = 0$, no biasing or rotation takes place. Estimating an optimum value of a that maximizes the simulation gain for given threshold η is an easy problem.

For the case $\mathbf{R} \neq \mathbf{I}$ we consider a unit vector $\mathbf{u}_1$ having elements denoted by t_i, $i = 1, \ldots, N$. By analogy we set

$$\mathbf{T} = \begin{pmatrix} t_1 - 1 & t_1 & . & . & t_1 \\ t_2 & t_2 - 1 & . & . & t_2 \\ . & & & & . \\ . & & & & . \\ t_N & . & . & . & t_N - 1 \end{pmatrix} \tag{3.16}$$

Again, when $a = 1$ the biased vector is just $\left(\sum_{i=1}^{N} x_{1i}\right)\mathbf{u}_1$ and is collinear with $\mathbf{u}_1$.

An interesting phenomenon takes place with this biasing scheme. The weighting function in (3.11) depends on the matrix $\mathbf{A}$ which in turn depends (apart from on the biasing parameter a) on the elements of the unit vector $\mathbf{u}_1$ through

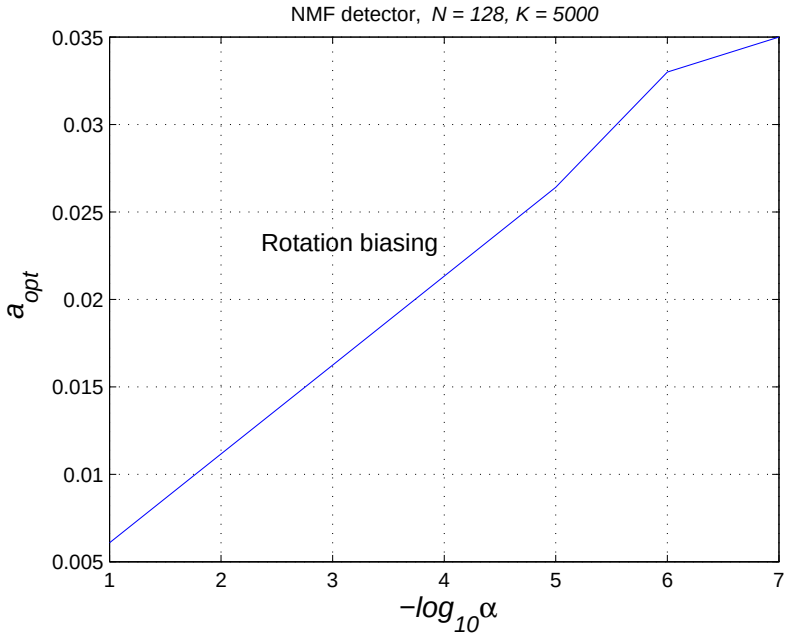

Figure 3.11: Optimum biasing parameter for the $\mathbf{T}$ matrix in rotation biasing, $\mathbf{R} = \mathbf{I}$.

the $\mathbf{T}$ matrix. The IS estimator in (3.12) estimates $\alpha_{\text{NMF}} = E\{1(\mathcal{A})\}$ which, from (3.10), is known to be independent of the particular unit vector $\mathbf{u}_1$ being used. Hence, if the estimator $\widehat{\alpha}_{\text{NMF}}$ is a reasonably good one, then it will be largely unaffected by the choice of the biasing matrix $\mathbf{A}$ (and hence of $\mathbf{T}$). However, the gain of the proposed IS scheme depends on the I-function given in (3.13) and this depends on the weighting function and therefore on the matrix $\mathbf{T}$. Thus the maximum gain for the optimized IS scheme will be affected by the choice of unit vector $\mathbf{u}_1$. Since $\mathbf{u}_1 = \mathbf{R}^{-1/2}\mathbf{s}/\|\mathbf{R}^{-1/2}\mathbf{s}\|$ it follows that the IS performance of this rotation biasing scheme depends on the data covariance matrix $\mathbf{R}$ in force; this is despite the fact that the detector FAP is independent of $\mathbf{R}$. Decoupling the biasing matrix $\mathbf{A}$ from $\mathbf{u}_1$ is of course not possible. Thus there may exist some $\mathbf{R}$ which gives a best IS simulation gain for given detector constants η and N. Furthermore, it may be possible to achieve some invariance of IS gain to covariance matrix $\mathbf{R}$ by using a matrix normalization for the $\mathbf{T}$ or $\mathbf{A}$ matrices. These issues are not investigated here and in the interests of expediency we present results obtained by simulating the $\mathbf{R} = \mathbf{I}$ case. We will visit rotation biasing later in this investigation.

The simulation results are in Figures 3.11 and 3.12 which show the estimated optimum biasing parameter and IS gains as functions of FAP respectively. As is evident from Fig 3.12 the simulation gain is not very high, being approximately 4×10^4 at a FAP of 10^{-6}.

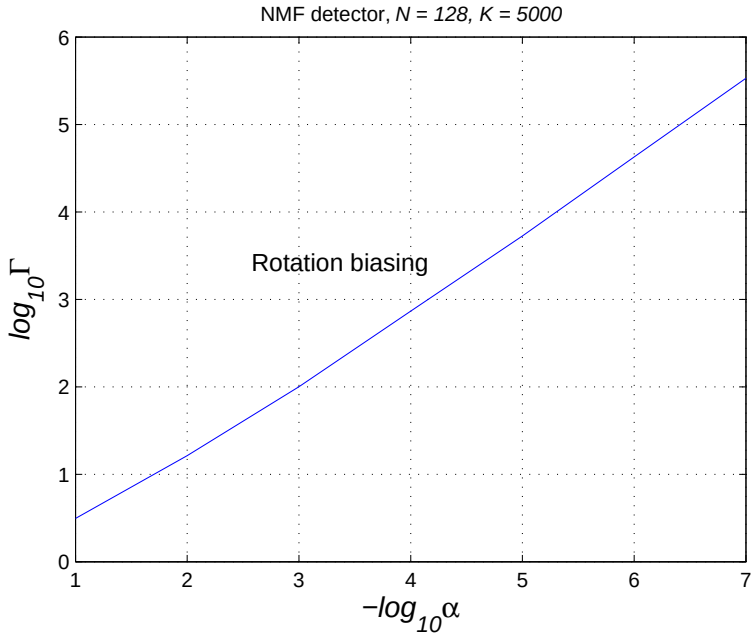

Figure 3.12: Simulation gain for rotation biasing, $\mathbf{R} = \mathbf{I}$.

3.3 The NAMF STAP detector

This section deals with the fast estimation of FAP for the NAMF detector. Developing an IS procedure for this is made easier by first setting up an alternate IS mechanization for FAP estimation of the AMF detector studied in Chapter 1.

3.3.1 FAP estimation for the AMF revisited

The alternate procedure for the AMF detector uses the g-method with two-dimensional biasing.

Reproducing (2.8) and (2.9) on page 20 we have

$$G = \frac{d}{L}\,\mathcal{P}_{AA}\,y \tag{3.17}$$

and

$$G(l) = \frac{d}{L}\,\mathcal{P}_{AA}\,y(l) \tag{3.18}$$

where

$$\begin{aligned} y &\equiv \mathbf{z}_A - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\mathbf{z}_B \\ y(l) &\equiv \mathbf{z}_A(l) - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\mathbf{z}_B(l) \end{aligned} \tag{3.19}$$

The AMF test of (1.3) on page 3 can therefore be written as

$$|y|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\eta}{L} \sum_{l=1}^{L} |y(l)|^2 \tag{3.20}$$

We know from Section 2.4 that, conditioned on $\mathbf{z}_B$ and $\{\mathbf{z}_B(l)\}_1^L$ (or, for short, the B-vectors), Y and $\{Y(l)\}_1^L$ are zero mean uncorrelated Gaussian random variables with variances and covariance as in (2.11), (2.12), and (2.13). From [2] it is known that

$$\sum_{l=1}^{L} |y(l)|^2 \overset{d}{=} \sum_{l=1}^{L-N+1} |w(l)|^2 \tag{3.21}$$

where the $w(l)$ are i.i.d. each with distribution $\mathcal{CN}_1(0,1)$. Moreover, $Y/M_B^{1/2}$ is, conditioned on the B-vectors, also distributed as $\mathcal{CN}_1(0,1)$ where

$$M_B \equiv 1 + \Sigma_B$$

and

$$\Sigma_B = \mathbf{z}_B^\dagger \left(\sum\nolimits_{l=1}^{L} \mathbf{z}_B(l) \mathbf{z}_B(l)^\dagger \right)^{-1} \mathbf{z}_B \tag{3.22}$$

The test in (3.20) then takes the form

$$u \underset{H_0}{\overset{H_1}{\gtrless}} \eta' \sum_{l=1}^{L-N+1} u(l) \tag{3.23}$$

where $U \equiv |Y|^2/M_B$ and $\{U(l) \equiv |w(l)|^2\}_1^{L-N+1}$ are all unit exponential and i.i.d., and $\eta' \equiv \eta/(LM_B)$. Once again, this is in the form of the usual CA-CFAR test when it is conditioned on the B-vectors. Hence the FAP of the AMF detector can be written as

$$\begin{aligned} \alpha_{\text{AMF}} &= P\Big(U \geq \eta' \sum\nolimits_{l=1}^{L-N+1} U(l)\Big) \\ &= E\Big\{ P\Big(U \geq \eta' \sum\nolimits_{l=1}^{L-N+1} U(l) \Big| \text{ B-vectors}\Big)\Big\} \\ &\triangleq E\{g(\Sigma_B)\} \end{aligned} \tag{3.24}$$

where

$$g(\Sigma_B) = \frac{1}{[1+\eta/(LM_B)]^{L-N+1}} \tag{3.25}$$

We can therefore estimate the FAP using the g-method with an IS simulation that biases the B-vectors. This estimator is

$$\widehat{\alpha}_{\text{AMF}} = \frac{1}{K} \sum_{1}^{K} g(\Sigma_B)\, W(\mathbf{z}_B, \mathbf{z}_{BL}); \quad \sim f_\star \tag{3.26}$$

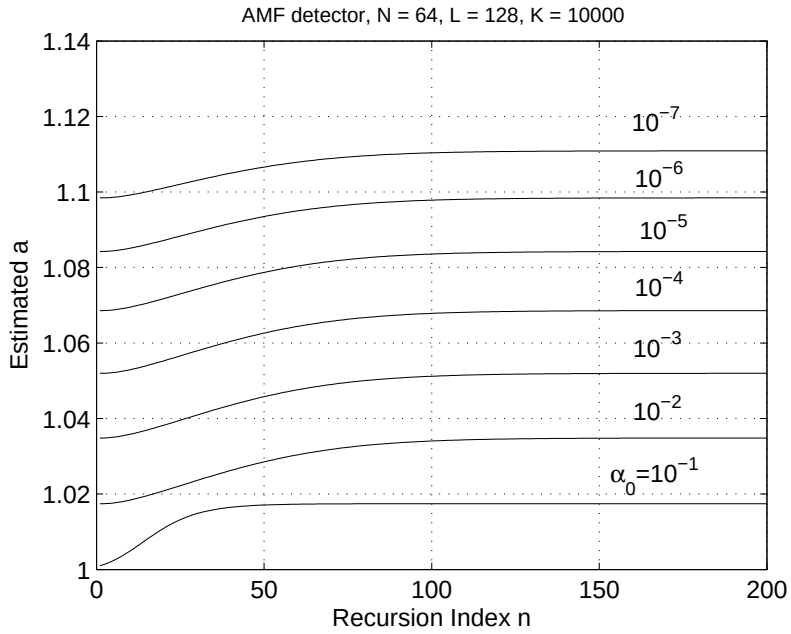

Figure 3.13: Optimum scaling for primary B-vectors.

where $\mathbf{z}_{BL} \equiv (\mathbf{z}_B(1), \ldots \mathbf{z}_B(L))'$.

Biasing of the B-vectors contained in Σ_B must produce an increase in the value of the g-function in (3.25). This means that Σ_B must be made to increase, which can easily be accomplished by scaling up the primary B-vector $\mathbf{z}_B$ and scaling down the secondary B-vectors $\mathbf{z}_B(l)$. A 2-d biasing scheme results, which needs to be optimized adaptively. As in Section 1.3.1 on page 3, the primary and secondary scaling parameters are chosen as $a^{1/2}$ and $\theta^{1/2}$ respectively, where $a \geq 1$ and $0 < \theta \leq 1$. The weighting function is easily shown to be

$$W(\mathbf{z}_B, \mathbf{z}_{BL}) = a^{N-1}\theta^{L(N-1)} \exp\big(-\mathbf{z}_B^\dagger \mathbf{z}_B(1-1/a)\big) \cdot \exp\Big(-(1-1/\theta)\sum_{l=1}^{L} \mathbf{z}_B(l)^\dagger \mathbf{z}_B(l)\Big) \quad (3.27)$$

The rest of the optimization procedure is as discussed in Section 1.3.

The results obtained from this 2-d IS simulation have been compared with the g-method used with input biasing, the technique which was developed in Chapter 1. Figures 3.13 - 3.16 contain these and are self-explanatory. From Fig 3.16 it is clear that applying the g-method with biasing of B-vectors is a more powerful IS scheme than the corresponding method used with input biasing of all secondary vectors, which was the subject of Chapter 1. This illustrates the points made in Section 3.1.

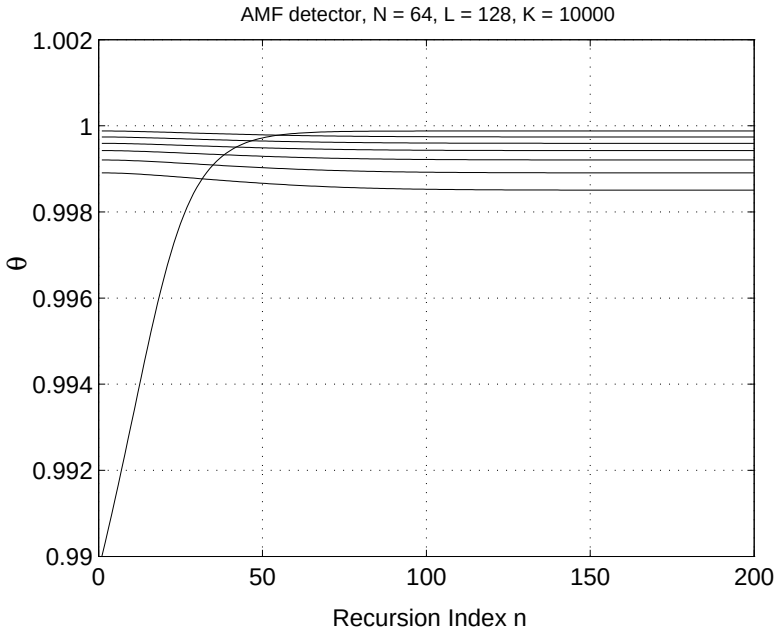

Figure 3.14: Optimum scaling for secondary B-vectors.

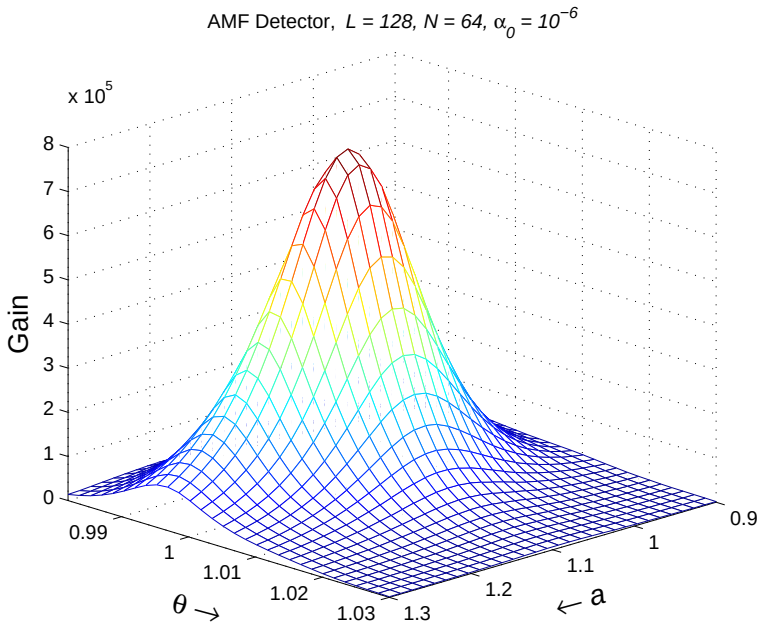

Figure 3.15: IS gain surface for 2-d biasing of B-vectors.

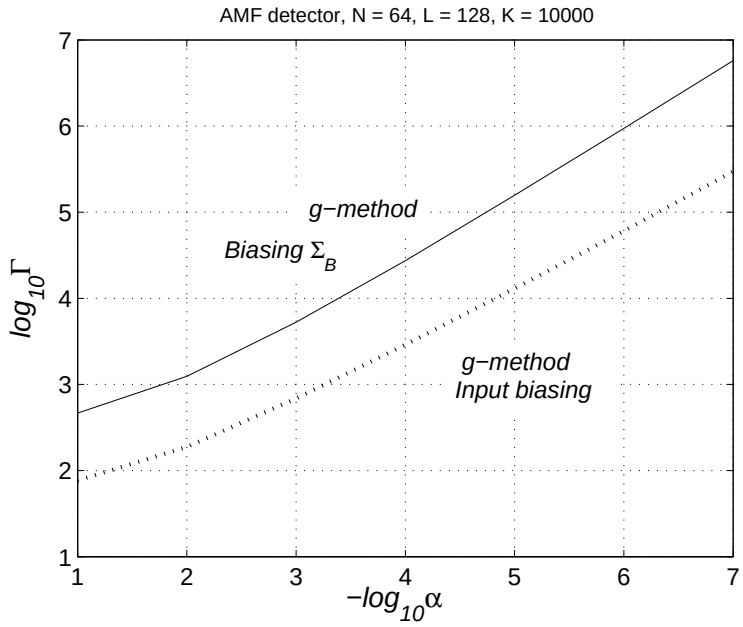

Figure 3.16: IS gain comparison for both g-methods.

3.3.2 IS for the NAMF detector

The NAMF detection statistic is given by

$$\frac{|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}|^2}{(\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{s})} \cdot \frac{1}{(\mathbf{x}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{3.28}$$

Using (3.23), this can be written as

$$\frac{u}{\sum_{l=1}^{L-N+1} u(l)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\eta}{LM_B} (\mathbf{x}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}) \tag{3.29}$$

With the transformations in Section 2.4 on page 19, the normalization term in the RHS of the above equation becomes

$$\mathbf{x}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x} = L\, \mathbf{z}^\dagger \mathcal{S}^{-1}\mathbf{z} \tag{3.30}$$

with $\mathcal{S}$ defined in (2.7) on page 20. From the discussion on pages 120 and 121 of [2], it turns out that the quantity $\mathbf{z}^\dagger \mathcal{S}^{-1}\mathbf{z}$, denoted therein as Σ, can be expressed as

$$\mathbf{z}^\dagger \mathcal{S}^{-1}\mathbf{z} \;=\; \frac{|y|^2}{\sum_{l=1}^{L} |y(l)|^2} + \Sigma_B$$

$$\stackrel{d}{\equiv} \quad \frac{|y|^2}{\sum_{l=1}^{L-N+1} |w(l)|^2} + \Sigma_B$$

$$= \quad M_B \cdot \frac{u}{\sum_{l=1}^{L-N+1} u(l)} + \Sigma_B \tag{3.31}$$

using the definitions of Section 3.3.1. Combining (3.29), (3.30), and (3.31) yields the test in the form

$$u \underset{H_0}{\overset{H_1}{\gtrless}} \eta_o \frac{\Sigma_B}{1+\Sigma_B} \sum_{l=1}^{L-N+1} u(l) \tag{3.32}$$

where $\eta_o = \eta/(1-\eta)$. Conditioned on the B-vectors this is again a CA-CFAR test resulting in the FAP

$$\alpha_{\text{NAMF}} = E\{g_N(\Sigma_B)\} \tag{3.33}$$

where

$$g_N(\Sigma_B) = \frac{1}{[1+\eta_o \Sigma_B/(1+\Sigma_B)]^{L-N+1}} \tag{3.34}$$

The FAP estimator is then

$$\widehat{\alpha}_{\text{NAMF}} = \frac{1}{K} \sum_{1}^{K} g_N(\Sigma_B)\, W(\mathbf{z}_B, \mathbf{z}_{BL}); \quad \sim f_\star \tag{3.35}$$

with the weighting function W being the same as in (3.27) excepting that the roles of the two biasing parameters are exchanged, that is, $0 < a \leq 1$ and $\theta \geq 1$. Simulation results are shown in Figures 3.17 - 3.22. IS results for threshold estimation have been compared with the analytical expression available in [15]. A quick way of determining the threshold for the NAMF detector is using the cubic interpolation polynomial :

$$\eta_{\text{NAMF}} = 7.96 \times 10^{-5} x^3 - 0.003116\, x^2 + 0.0723\, x + 7.77 \times 10^{-5} \tag{3.36}$$

where $x = -\log_{10} \alpha$.

3.4 Conclusion

In this chapter we have introduced and developed IS techniques suitable for analysis of the NMF class of STAP detectors. In particular, the method based on rotation of the primary data vector could be an effective solution for the difficult problem of characterizing the low-rank version of NAMF detectors. This is a matter for further investigation. Two dimensional biasing methods have been used to produce new IS results for the AMF and NAMF algorithms.

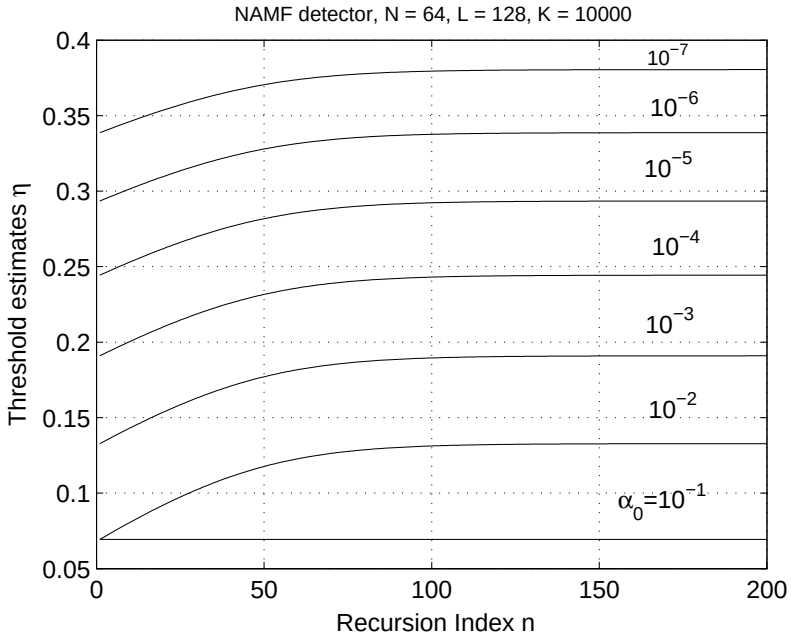

Figure 3.17: Thresholds through inverse IS.

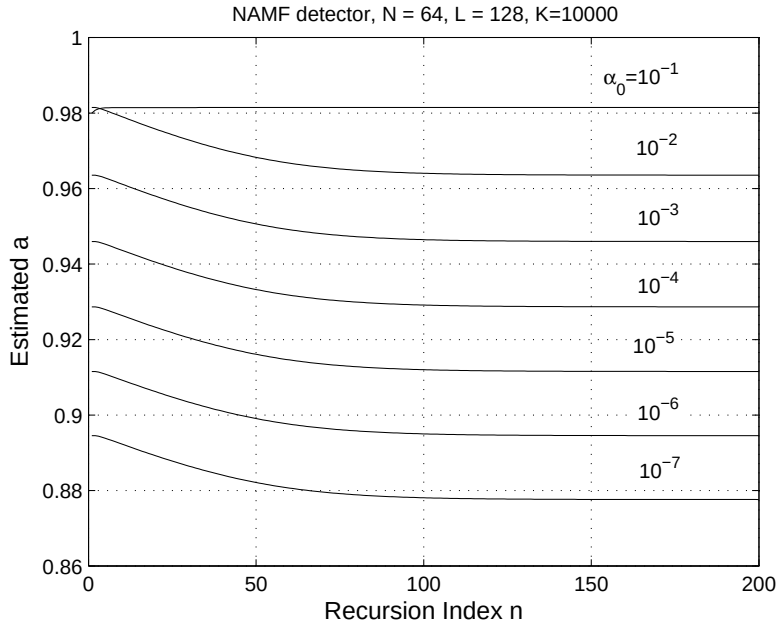

Figure 3.18: Optimum scaling for primary B-vectors.

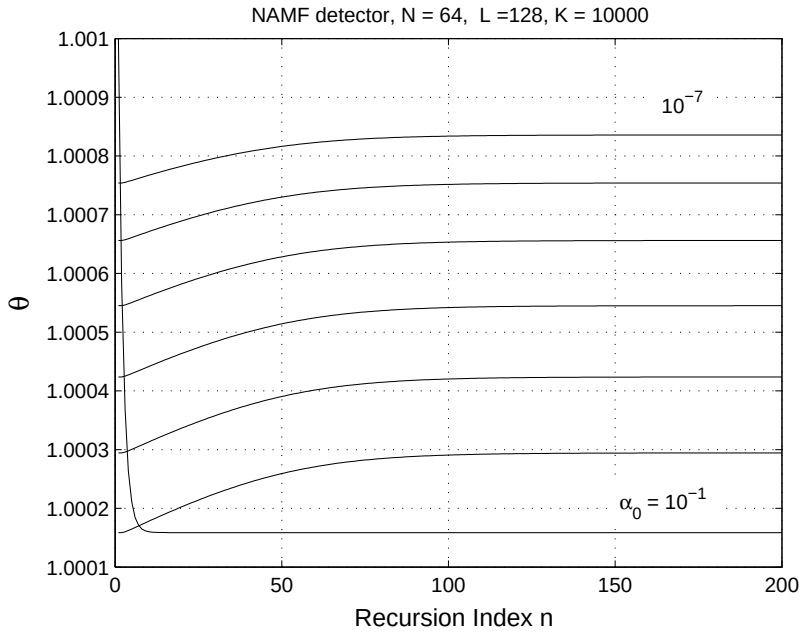

Figure 3.19: Optimum scaling for secondary B-vectors.

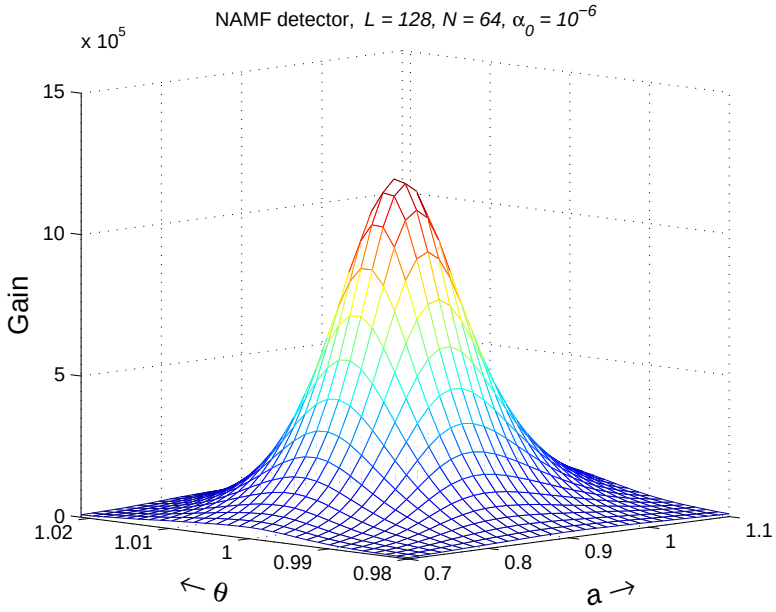

Figure 3.20: IS gain surface for 2-d biasing of B-vectors.

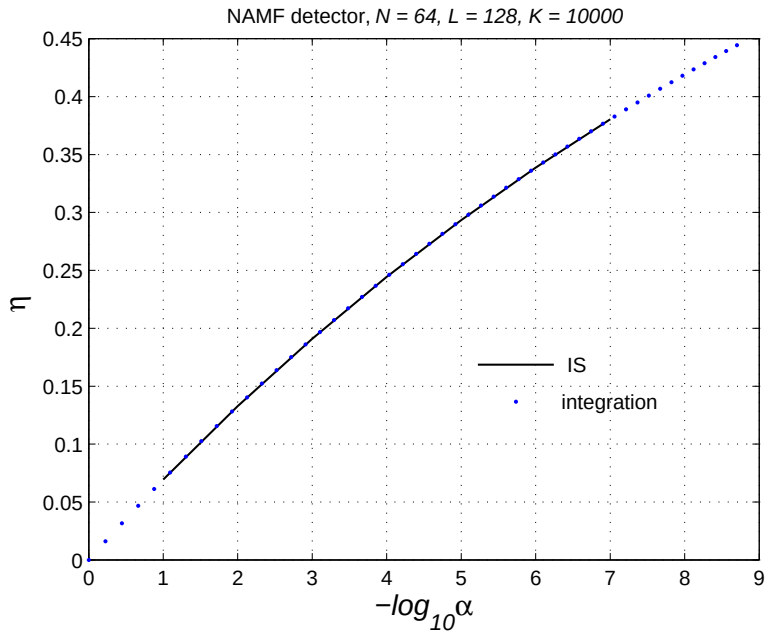

Figure 3.21: Comparison of simulation and numerical integration for thresholds.

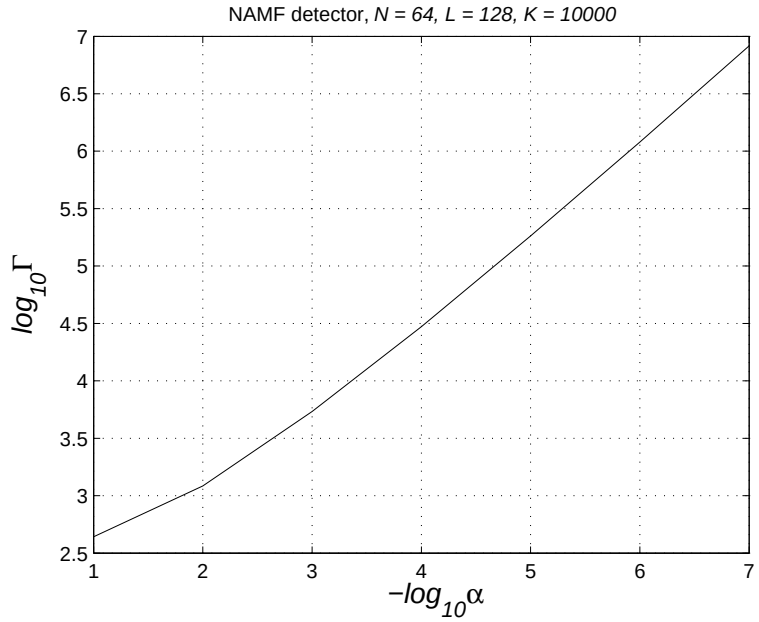

Figure 3.22: IS gain as function of FAP.

Chapter 4

Geometric mean and envelope-law NAMF STAP detectors

4.1 Introduction

Following the developments of Chapter 2, we propose two NAMF detector variants. These are the envelope and geometric-mean based versions. They are referred to hereafter as the E-NAMF detector and the GM-NAMF detector. In this chapter, their CFAR property is established and threshold settings for the detectors for specified false alarm probability (FAP) is accomplished using fast simulation based on importance sampling. The two variants show robustness in detection probability in the presence of interfering targets contaminating the training data while having zero loss in homogenous Gaussian interference compared to the usual square-law NAMF detector. This is shown for both fluctuating and non-fluctuating target models. Performance comparisons with the AMF variants are provided.

4.2 E-NAMF detector

This detector is constructed in the same way by which we defined the E-AMF detector. Consider the test

$$\frac{|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}|}{\sqrt{\mathbf{x}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}}} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\eta_e}{L}\sum_{l=1}^{L}|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}(l)| \tag{4.1}$$

which we refer to as the E-NAMF detector. As before, we expect this detector form to display some robustness against interfering targets in the secondary data compared to the usual NAMF detector.

4.2.1 Asymptotics

The asymptotic FAP behavior of this detector as $L \to \infty$ (or known covariance matrix $\mathbf{R}$) is obtained from

$$\begin{aligned} \alpha_{L\to\infty} &= P\left(\frac{|\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{X}|}{\sqrt{\mathbf{X}^\dagger \mathbf{R}^{-1}\mathbf{X}}} \geq \eta_e E\{|\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{X}(l)|\} \right) \\ &= P\left(\frac{|\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{X}|^2}{\mathbf{X}^\dagger \mathbf{R}^{-1}\mathbf{X}} \geq \frac{\eta_e^2 \pi}{4}\, \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s} \right) \\ &= \left(1 - \frac{\pi}{4}\,\eta_e^2\right)^{N-1} \end{aligned} \tag{4.2}$$

the first line above following from convergence arguments and the law of large numbers applied to the independent and identically distributed sequence $\{\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{X}(l)\}$, the second by noting that $\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{X}(l)$ is distributed as $\mathcal{CN}_1(0, \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s})$, and the third from the fact that the second line represents the FAP $(1-\nu)^{N-1}$ of a normalized matched filter (NMF) STAP detector (the statistic of which is shown in (3.1)) with threshold ν ([15]) where $\nu = \eta_e^2\pi/4$. The asymptotic threshold can be calculated from the above expression.

4.2.2 Alternate form and CFAR property

Using (3.18) and (3.19) of section 3.3.1, the E-NAMF test of (4.1) can be rewritten as

$$\frac{|y|}{\sqrt{\mathbf{x}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}}} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\eta_e}{L} \sum_{l=1}^{L} |y(l)| \tag{4.3}$$

Then, combining (3.30) and the first line of (3.31) and substituting for $\mathbf{x}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}$ in the above yields the test

$$\frac{|y|}{\sqrt{L\Sigma_B + L|y|^2 \Big/ \sum_{l=1}^{L} |y(l)|^2}} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\eta_e}{L} \sum_{l=1}^{L} |y(l)| \tag{4.4}$$

So if we replace y and $y(l)$ by G and $G(l)$ as defined in [10] respectively in the above, then the test remains unchanged. This implies, from the proposition of [10] and the arguments leading up to it, that the test is CFAR.

4.3 GM-NAMF detector

The geometric-mean version, referred to as the GM-NAMF detector, is defined as

$$\frac{|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}|}{\sqrt{\mathbf{x}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_g \left(\prod_{l=1}^{L} |\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}(l)| \right)^{1/L} \tag{4.5}$$

where η_g is the threshold. The square-law version is identical, being just a square of the above test expression.

4.3.1 Asymptotics

The asymptotic FAP for this detector is given by

$$\begin{aligned}\alpha_{L\to\infty} &= P\bigg(\log\frac{|\mathbf{s}^\dagger\mathbf{R}^{-1}\mathbf{X}|^2}{\mathbf{X}^\dagger\mathbf{R}^{-1}\mathbf{X}} \geq 2\log\eta_g \\ &\qquad\qquad + E\{\log|\mathbf{s}^\dagger\mathbf{R}^{-1}\mathbf{X}(l)|^2\}\bigg) \\ &= P\bigg(\log\frac{|\mathbf{s}^\dagger\mathbf{R}^{-1}\mathbf{X}|^2}{\mathbf{X}^\dagger\mathbf{R}^{-1}\mathbf{X}} \geq 2\log\eta_g \\ &\qquad\qquad - \gamma + \log(\mathbf{s}^\dagger\mathbf{R}^{-1}\mathbf{s})\bigg) \\ &= P\bigg(\frac{|\mathbf{s}^\dagger\mathbf{R}^{-1}\mathbf{X}|^2}{\mathbf{X}^\dagger\mathbf{R}^{-1}\mathbf{X}} \geq \eta_g^2\, e^{-\gamma}\,\mathbf{s}^\dagger\mathbf{R}^{-1}\mathbf{s}\bigg) \\ &= \left(1-\eta_g^2\, e^{-\gamma}\right)^{N-1} \end{aligned} \tag{4.6}$$

where γ is the Euler-Mascheroni constant.

4.3.2 Alternate form and CFAR property

Just as for the E-NAMF detector, the GM-NAMF test of (4.5) can be rewritten as

$$\frac{|y|}{\sqrt{\mathbf{x}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{x}}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_g\bigg(\prod_{l=1}^{L}|y(l)|\bigg)^{1/L} \tag{4.7}$$

and then as

$$\frac{|y|}{\sqrt{L\Sigma_B + L|y|^2\Big/\sum_{l=1}^{L}|y(l)|^2}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_g\bigg(\prod_{l=1}^{L}|y(l)|\bigg)^{1/L} \tag{4.8}$$

For the same reasons as for the E-NAMF detector, the GM-NAMF test is also CFAR.

4.4 FAP estimation

With some straightforward manipulations, the E-NAMF test in (4.4) can be rewritten as

$$|y|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\eta_e^2\Sigma_B}{LC_e}\left(\sum\nolimits_{l=1}^{L}|y(l)|\right)^2 \tag{4.9}$$

where

$$C_e \equiv 1 - \frac{\eta_e^2}{L}\frac{\left(\sum_{l=1}^{L}|y(l)|\right)^2}{\sum_{l=1}^{L}|y(l)|^2} \tag{4.10}$$

and the GM-NAMF in (4.5) as

$$|y|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\eta_g^2 L \Sigma_{_B}}{C_g} \left(\prod_{l=1}^{L} |y(l)|^2 \right)^{1/L} \tag{4.11}$$

where

$$C_g \equiv 1 - L\eta_g^2 \frac{\left(\prod_{l=1}^{L} |y(l)|^2 \right)^{1/L}}{\sum_{l=1}^{L} |y(l)|^2} \tag{4.12}$$

In writing (4.9) and (4.11) we have assumed that $C_e > 0$ and $C_g > 0$. This will be guaranteed if $\eta_e \leq 1$ and $\eta_g \leq 1$, respectively, and can be seen by an application of Hölder's inequality to the sums in (4.10) and using the fact that geometric means are smaller than arithmetic means in (4.12). As shown by simulation results, a wide range of FAPs is achieved for values of both thresholds less than unity and therefore we assume this restriction to hold. For brevity we denote the RHSs of both (4.9) and (4.11) by $\Sigma_{_B} D$, where D is a function of $\begin{pmatrix} y(1) & \cdots & y(L) \end{pmatrix} \triangleq \mathbf{y}_L^T$ and corresponding constants that depend on the detector in question. With this (generic) notation, the FAP can be written as

$$\begin{aligned}
\alpha &= P(|Y|^2 \geq \Sigma_{_B} D) \\
&= E_B\{P(|Y|^2 \geq \Sigma_{_B} D \,|\, \text{B-vec})\} \\
&= E_B\{E\{P(|Y|^2 \geq \Sigma_{_B} D \,|\, \text{B-vec}, \mathbf{Y}_L)\}\} \\
&= E_B\{E\{e^{-\Sigma_{_B} D/(1+\Sigma_{_B})} \,|\, \text{B-vec})\}\} \\
&= E_B\big\{E\big\{g(\Sigma_{_B}, \mathbf{Y}_L)\big|\, \text{B-vec}\big\}\big\}
\end{aligned} \tag{4.13}$$

where

$$g(\Sigma_{_B}, \mathbf{Y}_L) \equiv e^{-\Sigma_{_B} D/(1+\Sigma_{_B})} \tag{4.14}$$

and E_B denotes expectation over the distribution of the B vectors and E the expectation over the conditional distribution of $\mathbf{Y}_L$. The fourth line above follows because conditioned on the B vectors, Y and $\mathbf{Y}_L$ are independent and $|Y|^2/(1+\Sigma_{_B})$ is unit exponential.

From (2.7) on page 20, y and $y(l)$ are

$$\begin{aligned}
y &\equiv \mathbf{z}_{_A} - \sum\nolimits_{l=1}^{L} \mathbf{z}_{_A}(l)\mathbf{z}_{_B}(l)^\dagger \mathcal{S}_{BB}^{-1} \mathbf{z}_{_B} \\
y(l) &\equiv \mathbf{z}_{_A}(l) - \sum\nolimits_{i=1}^{L} \mathbf{z}_{_A}(i)\mathbf{z}_{_B}(i)^\dagger \mathcal{S}_{BB}^{-1} \mathbf{z}_{_B}(l)
\end{aligned} \tag{4.15}$$

Then the FAP in (4.13) can be further written as

$$\alpha = E_B\big\{E\big\{g(\Sigma_{_B}, \mathbf{Y}_L)\big|\, \text{B-vec}\big\}\big\}$$

$$
\begin{aligned}
&= \int\int g(\Sigma_B, \mathbf{Y}_L)\, f(\mathbf{Y}_L \big|\, \text{B-vec})\, f(\text{B-vec})\, d\mathbf{Y}_L\, d\text{B-vec} \\
&= \int\int g(\Sigma_B, \mathbf{Y}_L)\, f(\mathbf{Y}_L, \text{B-vec})\, d\mathbf{Y}_L\, d\text{B-vec} \\
&= E\{g(\Sigma_B, \mathbf{Y}_L)\} \\
&= E\{g(\Sigma_B, Y(1), \ldots, Y(L))\} \\
&= E\{g(\Sigma_B, Z_A(1), \ldots, Z_A(L))\} \\
&\triangleq E\{g(\Sigma_B, \mathbf{Z}_{AL})\} \\
&= \int\int g(\Sigma_B, \mathbf{z}_{AL})\, f(z_{AL}, \text{B-vec})\, d\mathbf{z}_{AL}\, d\text{B-vec} \\
&= \int\int g(\Sigma_B, \mathbf{z}_{AL})\, f(\mathbf{z}_{AL})\, f(\text{B-vec})\, d\mathbf{z}_{AL}\, d\text{B-vec} \\
&= \int\int g(\Sigma_B, \mathbf{z}_{AL})\, \frac{f(\mathbf{z}_{AL})}{f_\star(\mathbf{z}_{AL})}\, f_\star(\mathbf{z}_{AL})\, f(\text{B-vec})\, d\mathbf{z}_{AL}\, d\text{B-vec} \\
&\triangleq \int\int g(\Sigma_B, \mathbf{z}_{AL})\, W(\mathbf{z}_{AL})\, f_\star(\mathbf{z}_{AL})\, f(\text{B-vec})\, d\mathbf{z}_{AL}\, d\text{B-vec} \\
&= E_\star\{g(\Sigma_B, \mathbf{Z}_{AL})\, W(\mathbf{Z}_{AL})\} \qquad (4.16)
\end{aligned}
$$

where $g(\Sigma_B, \mathbf{Z}_{AL}) = g(\Sigma_B, \mathbf{Y}_L)$ and is given by (4.14). We perform IS only on the variables in $\mathbf{Z}_{AL} = (Z_A(1) \cdots Z_A(L))^T$, for simplicity. Noting that $\mathbf{Z}_{AL} \sim \mathcal{CN}_L(0, \mathbf{I})$, scaling down $\mathbf{Z}_{AL}$ with parameter θ leads to the weighting function

$$
W(\mathbf{z}_{AL}) = \theta^L\, e^{-\sum_1^L z_A^\dagger(l) z_A(l)(1-1/\theta)} \qquad (4.17)
$$

The IS estimator for FAP is therefore

$$
\widehat{\alpha} = \frac{1}{K} \sum_{i=1}^{K} [g(\Sigma_B, \mathbf{Z}_{AL})\, W(\mathbf{z}_{AL})]^{(i)}; \quad \mathbf{Z}_{AL} \sim f_\star(\mathbf{z}_{AL}), \text{B-vec} \sim f(\text{B-vec}) \qquad (4.18)
$$

4.5 Simulation results

The adaptive IS implementation is as usual. Simulation results for the FAP range $10^{-1} - 10^{-7}$ are shown in Figures 4.1 - 4.14, for both ENAMF and GNAMF detectors. The parameters used in the simulations are $L = 128$, $N = 64$ and the number of IS trials is $K = 50000$. Moreover, for $L = 128$, $N = 64$, thresholds for these detectors can be determined using the cubic interpolation

$$
\eta_e = 9.704 \times 10^{-4} x^3 - 1.85 \times 10^{-2} x^2 + 0.1595x + 0.15613 \qquad (4.19)
$$

$$\eta_g = 1.1439 \times 10^{-3} x^3 - 2.148 \times 10^{-2} x^2 + 0.188x + 0.184 \tag{4.20}$$

where $x = -\log_{10} \alpha$.

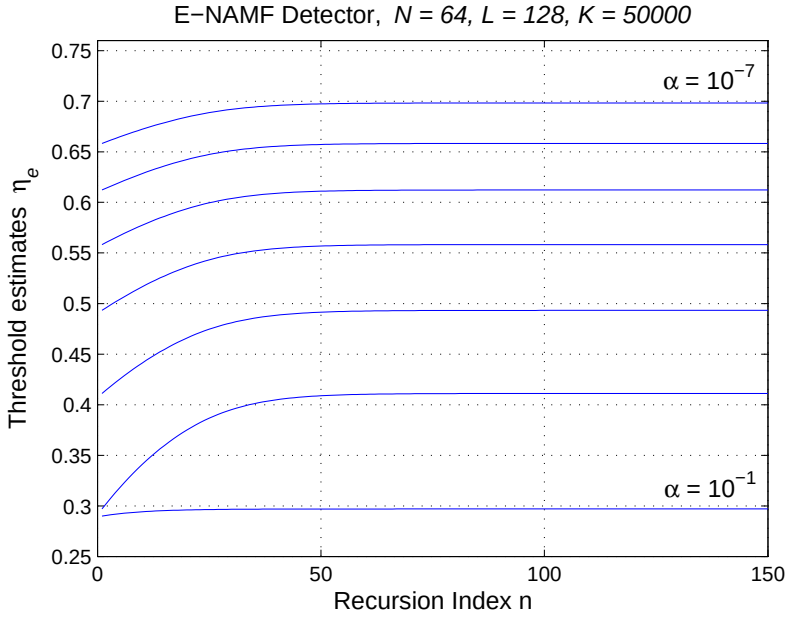

Figure 4.1: Inverse g-method thresholds.

4.5.1 P_D in homogeneous case

In this subsection we present the detection probability of the E-NAMF and GM-NAMF detectors in homogeneous background and compare it with that of the NAMF detector. We set FAP to 10^{-6}, the value of L and N are the same as in FAP simulations. All the simulations concerning detection probability are done using Monte Carlo (MC) technique and the number of trials used is 100000. The detection probability in case of non-fluctuating target can be seen in Figure 4.15. In this figure is also plotted the detection probability for the NAMF detector obtained using the analytical formula of [15]. The figure shows that the three detectors perform almost the same in the presence of homogeneous clutter. In fact, the E-NAMF and NAMF perform exactly the same while the GM-NAMF has a very small loss compared to the other two detectors. Simulations are also done to estimate detection probability when the Swerling I fluctuating target model is used. The result is presented in Figure 4.16. Similarly, in this case the detectors have almost the same performance. From these simulation results we can conclude that, in homogeneous case, regardless of the target fluctuation,

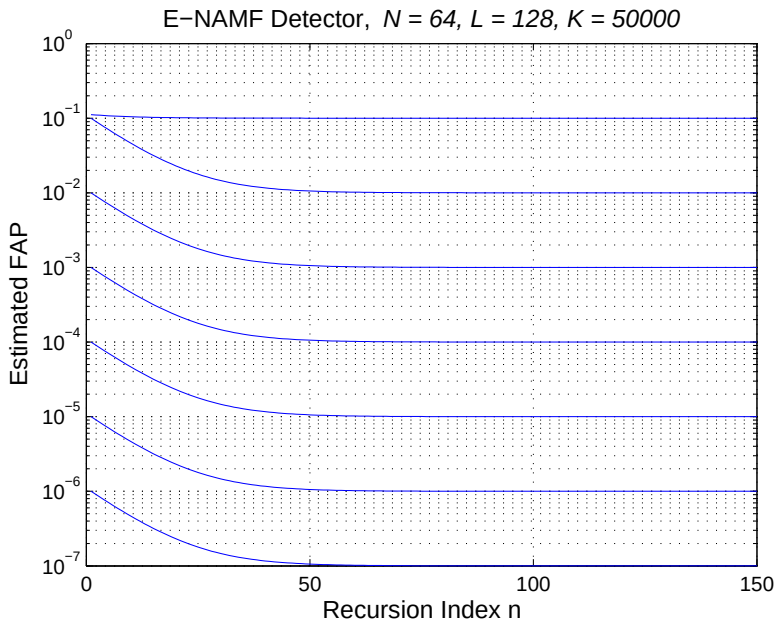

Figure 4.2: FAP estimates.

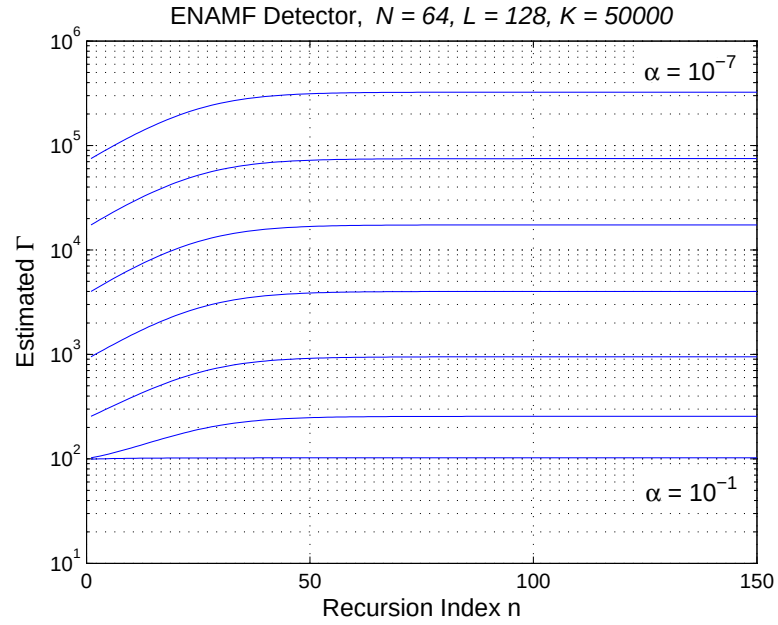

Figure 4.3: Gain estimates.

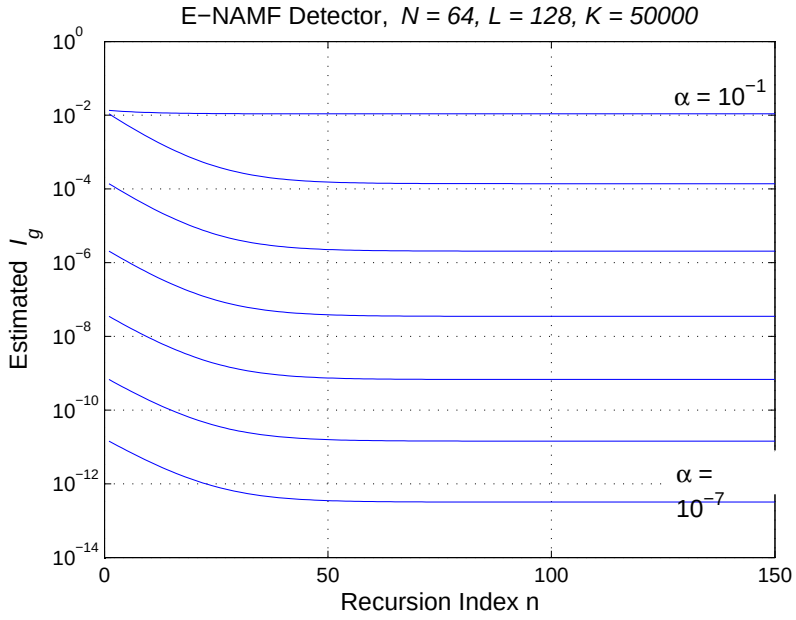

Figure 4.4: I_g function estimates.

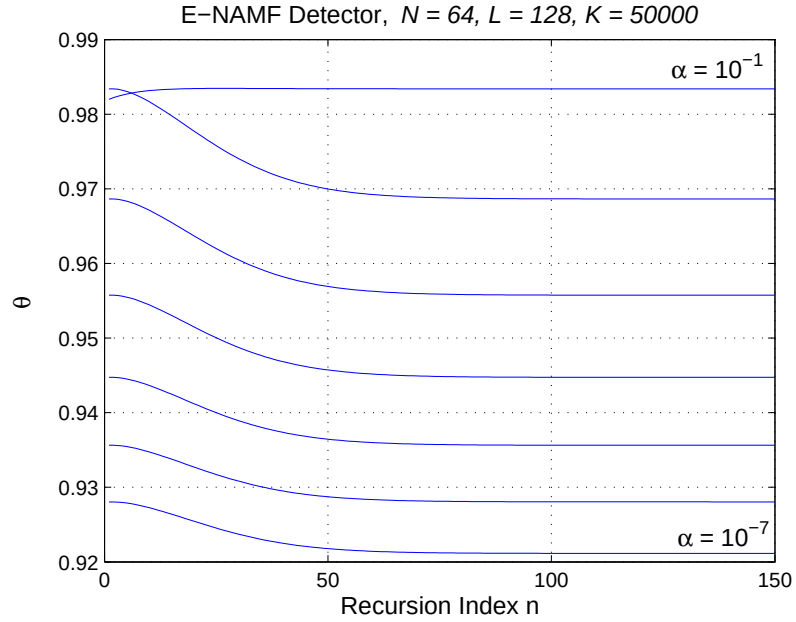

Figure 4.5: Optimum biasing parameter estimates.

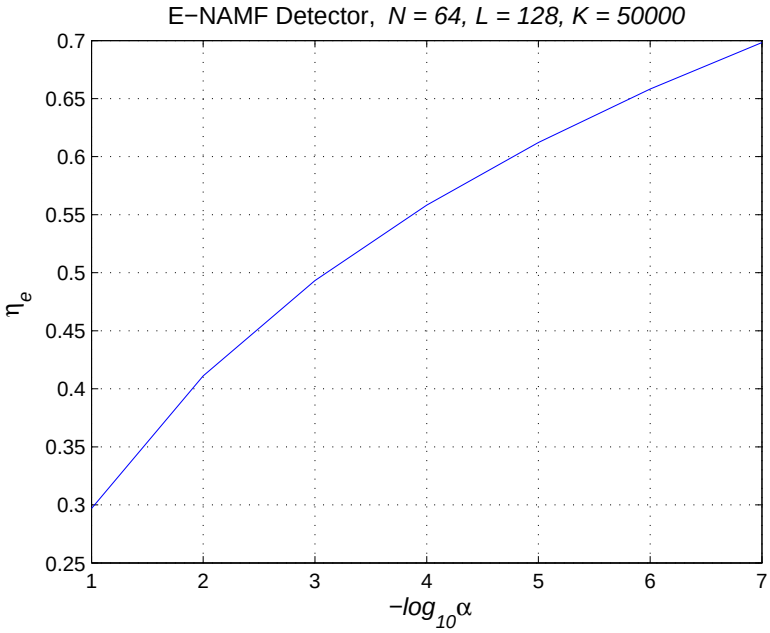

Figure 4.6: Threshold as function of FAP.

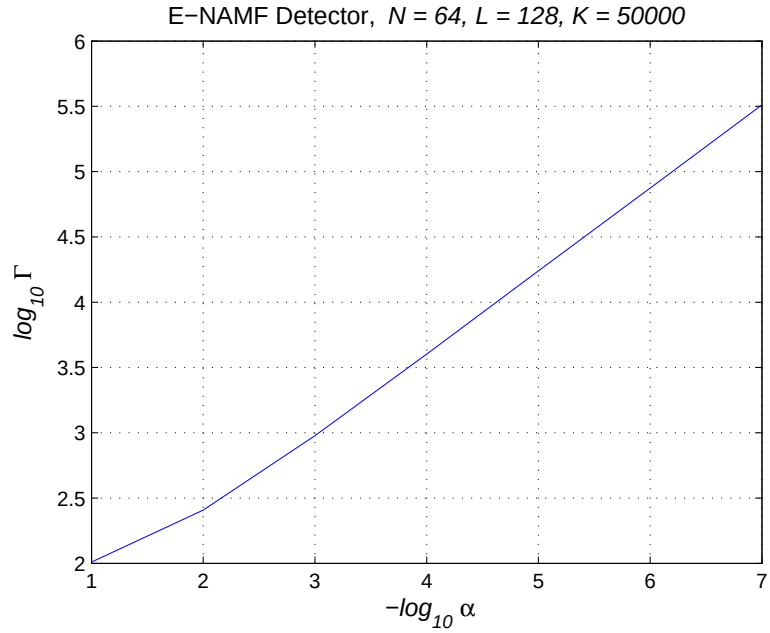

Figure 4.7: Gain as function of FAP.

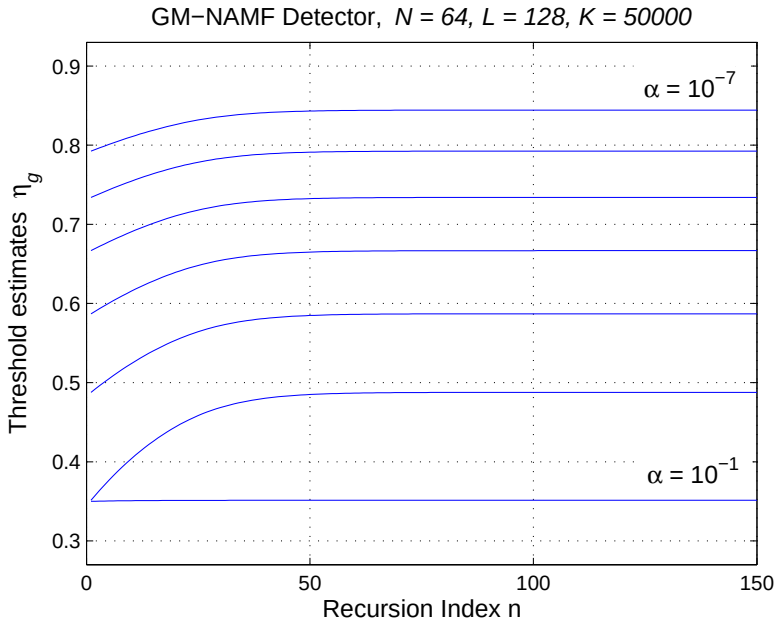

Figure 4.8: Inverse g-method thresholds.

both the E-NAMF detector and the GM-NAMF detector perform as well as the NAMF detector.

4.5.2 P_D in the presence of nonhomogeneities

In this subsection we will show the simulation results that compare the performance of the envelope, geometric mean and square law NAMF detectors in the presence of nonhomogeneities. The nonhomogeneities considered here are two interfering targets in the secondary data. In the simulations we assumed the interfering targets to have the same power and the same steering vector as the actual target. Figure 4.17 shows the comparison of the performance of E-NAMF, GM-NAMF and NAMF in the presence of 2 interferers when the target signal is constant, i.e. non-fluctuating. In Figure 4.18 is plotted the detection probability when both the target and the interferers are assumed to be fluctuating. For both target models, it can be seen that the NAMF detector, for a fix P_D, has a significant detection loss compared to the envelope and geometric mean variants and that the GM-NAMF performs slightly better than the E-NAMF detector. For example, at $P_D = 0.5$, the NAMF has detection loss about 1.2 dB w.r.t E-NAMF and 1.38 dB w.r.t GM-NAMF for the case of non-fluctuating target and about 2 dB w.r.t E-NAMF and 2.38 dB w.r.t GM-NAMF for the fluctuating model. This detection loss will increase for higher P_D. For example, in the case of Swerling 0 target model at $P_D = 0.9$, NAMF has about 3 dB loss

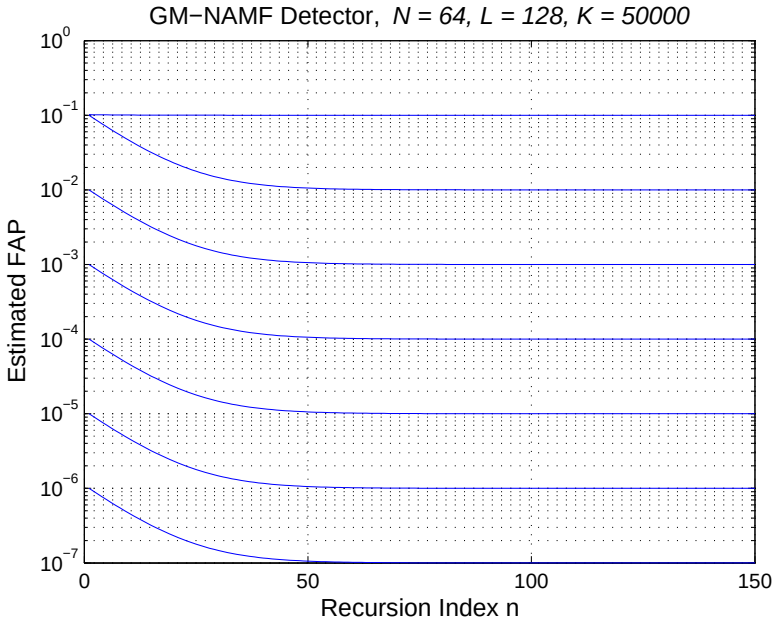

Figure 4.9: FAP estimates.

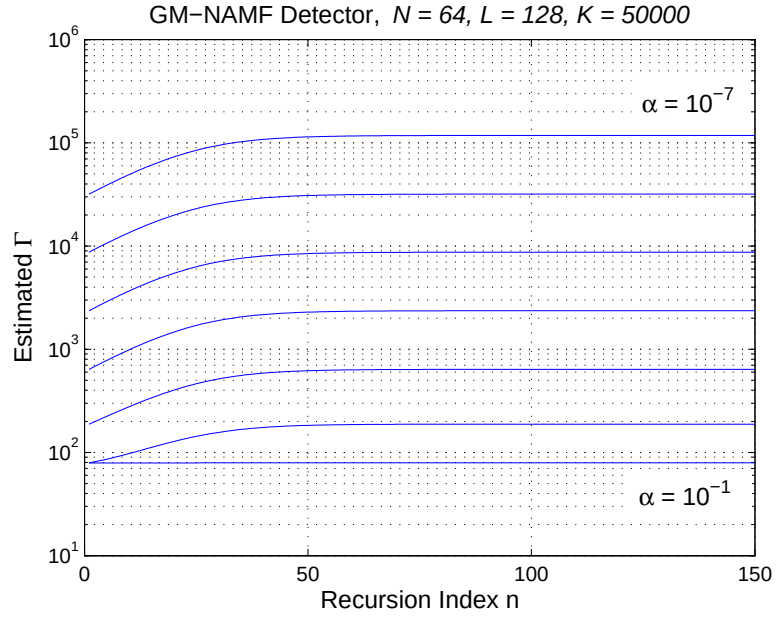

Figure 4.10: Gain estimates.

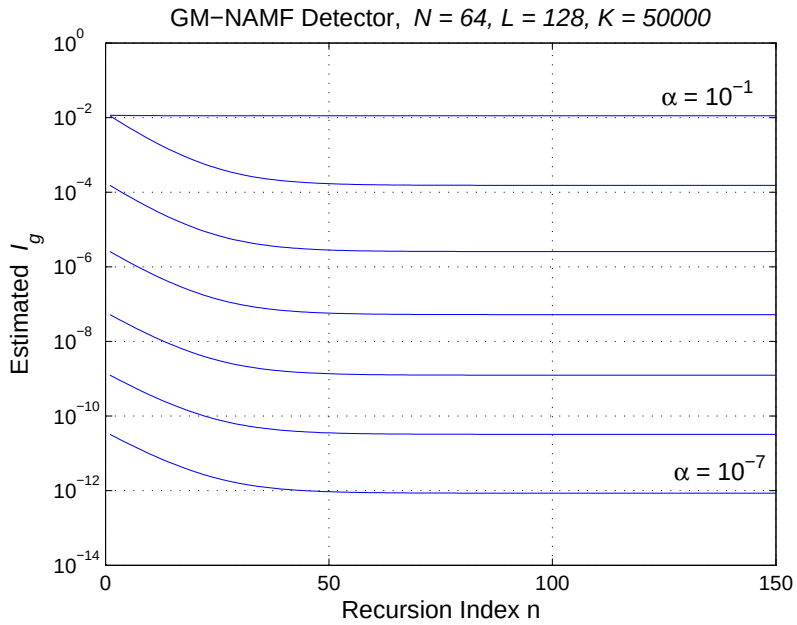

Figure 4.11: I_g function estimates.

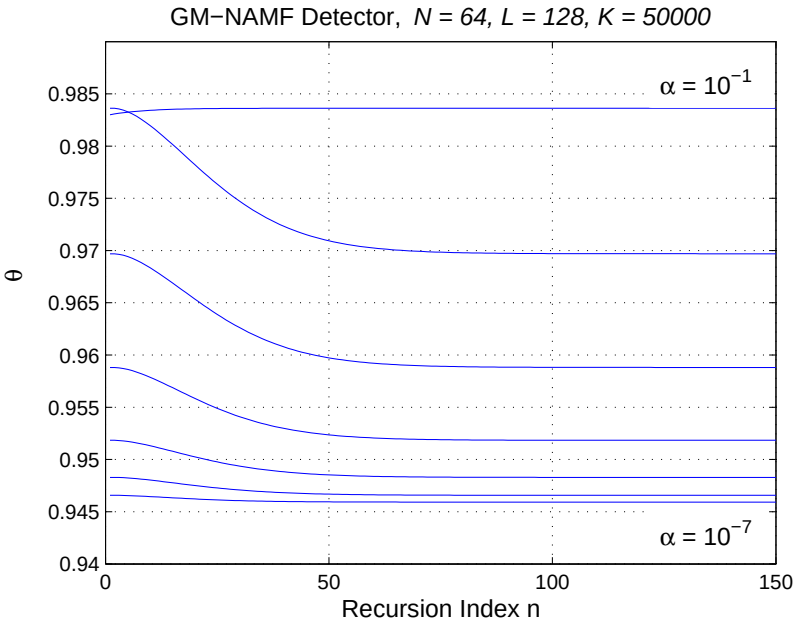

Figure 4.12: Optimum biasing parameter estimates.

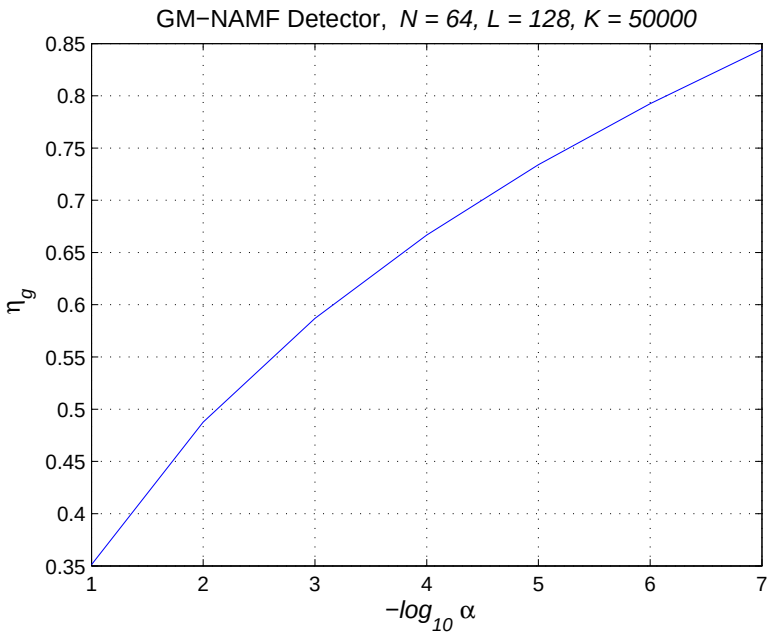

Figure 4.13: Threshold as function of FAP.

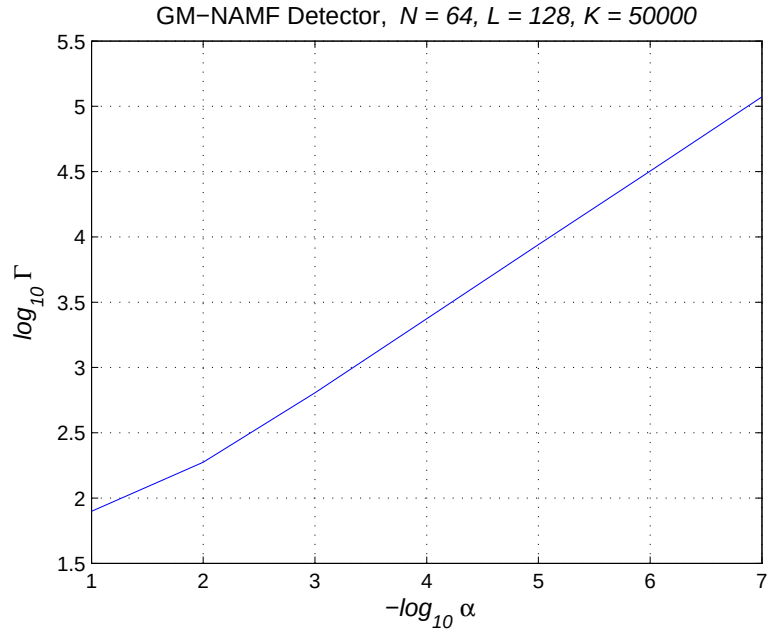

Figure 4.14: Gain as function of FAP.

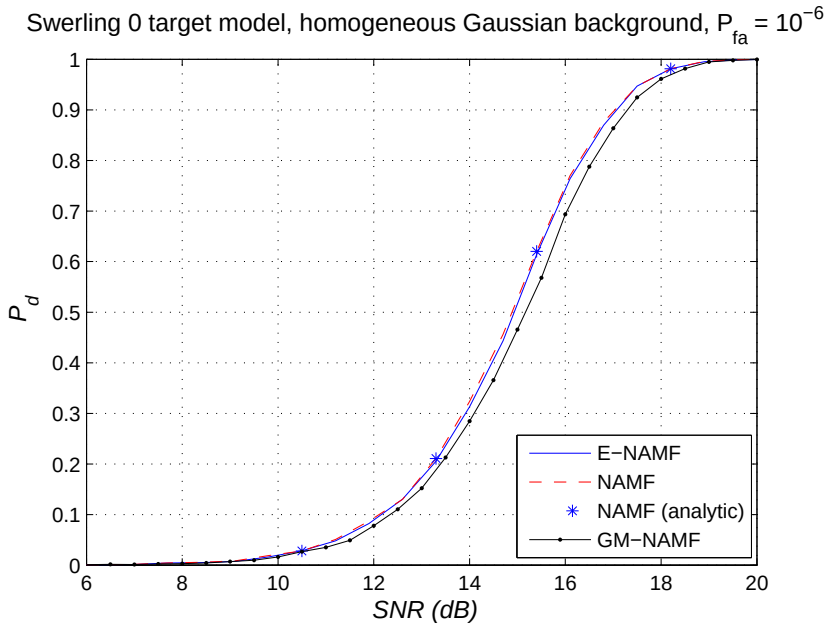

Figure 4.15: P_D in homogeneous case, Swerling 0 target.

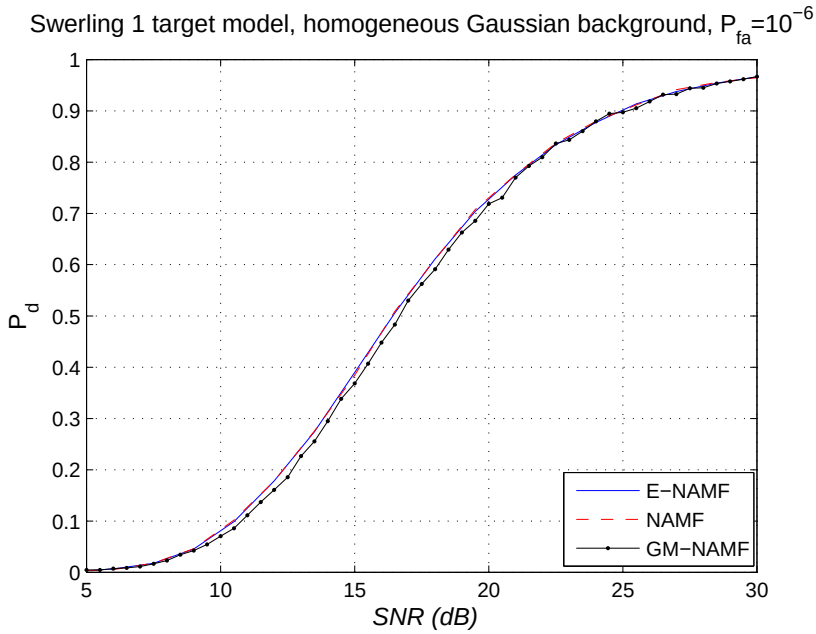

Figure 4.16: P_D in homogeneous case, Swerling I target.

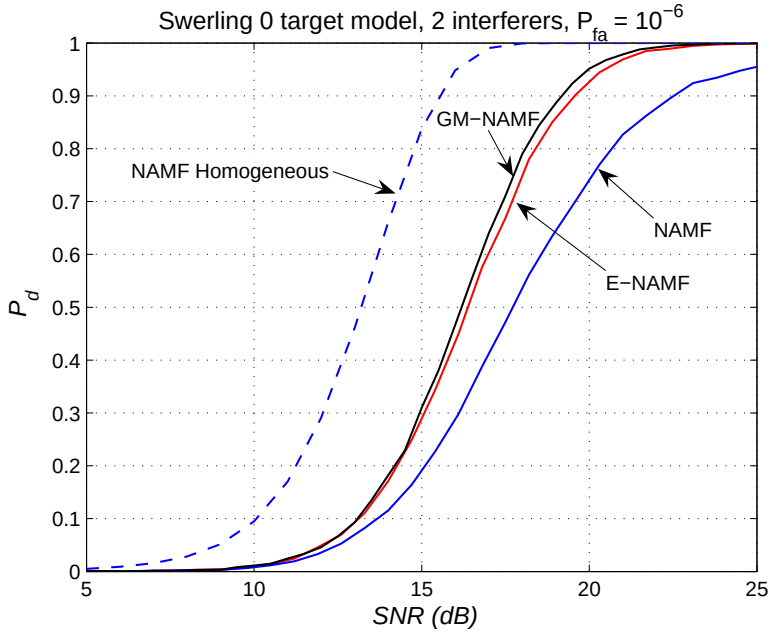

Figure 4.17: P_D in the presence of two interferers. Swerling 0 target.

w.r.t E-NAMF for the same case and 5.5 dB loss w.r.t NAMF in homogeneous Gaussian background. It has also to be noticed that in the fluctuating target scenario the presence of nonhomogeneities produces a loss in term of P_D performance much more significant that in the nonfluctuating case, especially in the region of high SNR. From these results we conclude that in the presence of nonhomogeneities the E-NAMF and GM-NAMF detectors are much robust than the NAMF detector. Also, we compare the detection performance of the class of NAMF detectors with the AMF detector and its envelope and geometric mean variants. In Figure 4.19 and 4.20 these performance are shown respectively for the Swerling 0 and Swerling 1 target model. The nonhomogeneities considered here are the same as in the previous case. Properties of the two detector classes are summarized in Table 4.1. From these results, we conclude that the geometric mean and envelope detectors of the NAMF class have the best detection performance in the presence of interfering targets.

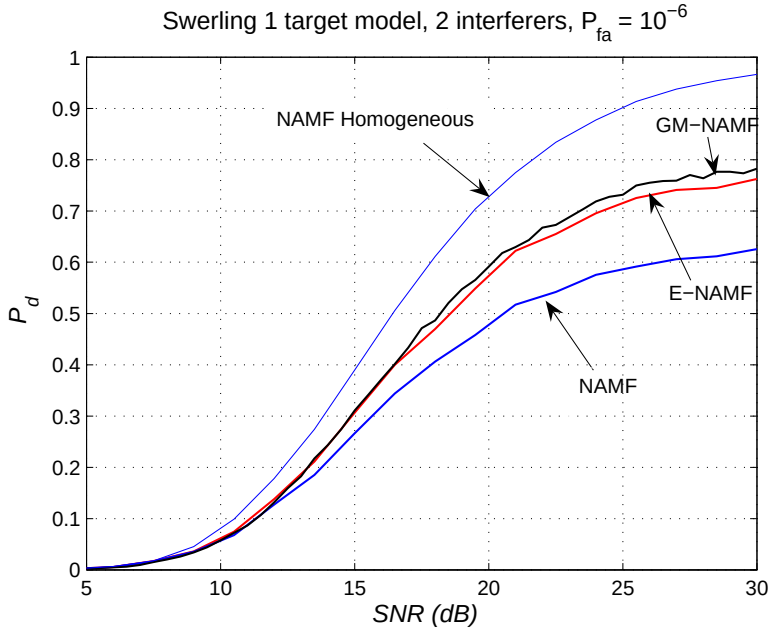

Figure 4.18: P_D in the presence of two interferers. Swerling 1 target.

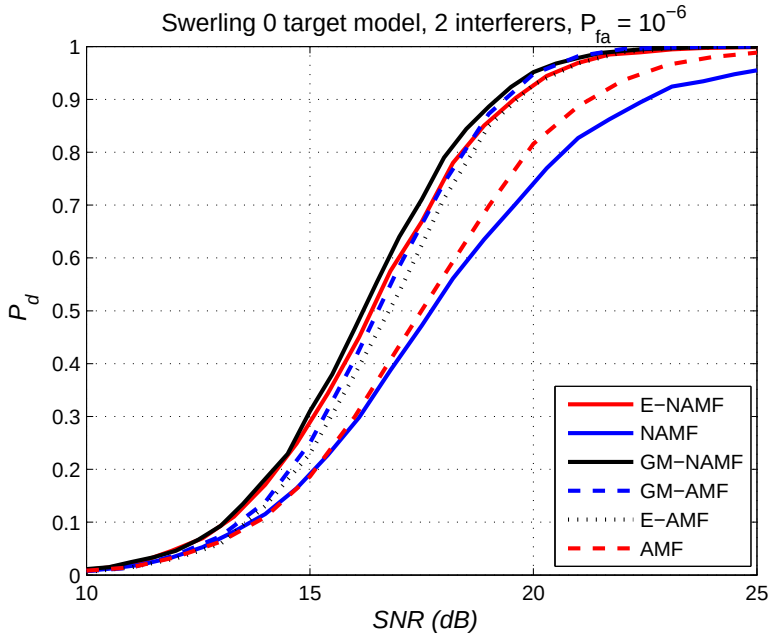

Figure 4.19: P_D comparison in the presence of two interferers. Swerling 0 target.

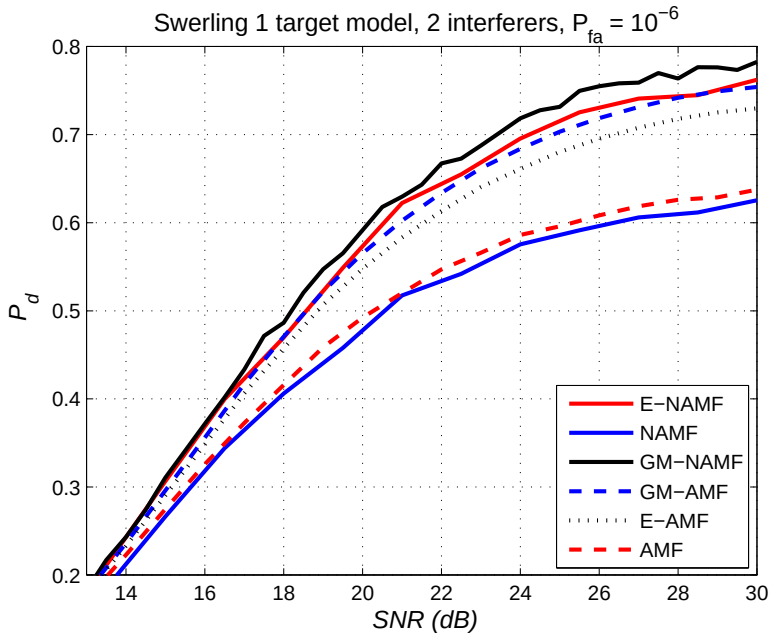

Figure 4.20: P_D comparison in the presence of two interferers. Swerling 1 target.

Table 4.1: P_D comparison of AMF and NAMF STAP detectors.

Q	AMF	E-AMF	GM-AMF	NAMF	E-NAMF	GM-NAMF
CFAR?	Yes	Yes	Yes	Yes	Yes	Yes
homog. P_d?	-	ϵ-loss	ϵ-loss	-	≈ 0 loss	≈ 0 loss
Robust?	No	Yes	Yes	No	Yes	Yes

4.6 Conclusion

Envelope-law and geometric-mean variants of the NAMF detector have been introduced. Their respective thresholds have been determined for a range of FAPs. The detectors have been shown to have better performance in the presence of outliers in the training data while maintaining almost equal performance as the standard square-law version in homogeneous Gaussian interference. Together with the variants proposed in Chapter 2, these detectors represent robust alternatives to conventional square-law processing.

Chapter 5

Low-Rank STAP detectors

5.1 Introduction

In this chapter the STAP detector based on the low-rank approximation of the normalized adaptive matched filter (LRNAMF) is investigated for its performance. Being computationally efficient, the LRNAMF detector is a candidate for future implementation in STAP radar detection. Its FAP performance is analytically investigated here.

In the second section the low-rank NMF detector is described. Subject to an approximation for the disturbance covariance matrix in a clutter dominated scenario, the FAP of the LRNMF detector is known via a simple formula, [16] and [17]. A brief description of the detector is reported here in order to summarize the low-rank approximation. The third section provides analytical derivation of the exact FAP of the LRNMF detector for data possessing an arbitrary covariance matrix. The fourth section discusses a nominal covariance model for simulation and gives further analytical results. The fifth section discusses the LRNAMF detector and shows how the analytical results for the LRNMF detector can be used to predict its performance. The sixth section provides simulation results followed by a conclusion.

5.2 The LRNMF detector

The low-rank approximation to the normalized matched filter is used for target detection in heterogeneous clutter scenarios. The background is assumed to consist of clutter plus white Gaussian noise. The binary hypothesis test can be written as

$$\begin{aligned} \mathrm{H}_0 : \mathbf{x} &= \mathbf{d} = \mathbf{c} + \mathbf{n} \\ \mathrm{H}_1 : \mathbf{x} &= a\mathbf{s} + \mathbf{d} = a\mathbf{s} + \mathbf{c} + \mathbf{n} \end{aligned} \tag{5.1}$$

where $\mathbf{x}$ is the primary data vector, $\mathbf{c}$ is the Gaussian clutter vector with covariance matrix $s\mathbf{R}_c$ with unknown level s and known structure, $\mathbf{n}$ denotes the additive white Gaussian noise vector with covariance matrix $\sigma^2\mathbf{I}$, where $\mathbf{I}$ is the $N \times N$ identity matrix and the noise power σ^2 is unknown, $\mathbf{s}$ is the steering vector and a is the unknown target amplitude. The vector $\mathbf{d}$ is used to represent the sum of the clutter and the white Gaussian noise. The covariance matrix of the disturbance $\mathbf{d}$ is

$$\mathbf{R}_d = s\mathbf{R}_c + \sigma^2\mathbf{I} \tag{5.2}$$

In many real cases the clutter covariance matrix $\mathbf{R}_c$ has rank r which is less than N. This fact will be used to approximate the inverse of the disturbance covariance matrix which will be used in the NMF test. The matrix $\mathbf{R}_d$ can be expressed as

$$\mathbf{R}_d = \mathbf{U}\mathbf{D}\mathbf{U}^\dagger \tag{5.3}$$

where $\mathbf{U}$ is the matrix whose columns are the normalized eigenvectors of $\mathbf{R}_d$ and $\mathbf{D}$ is the diagonal matrix of the eigenvalues of $\mathbf{R}_d$. When $\mathbf{R}_c$ has rank $r \ll N$, then $\mathbf{R}_d$ can be rewritten as

$$\mathbf{R}_d = \sum_{i=1}^{r}(s\lambda_i + \sigma^2)\mathbf{u}_i\mathbf{u}_i^\dagger + \sum_{i=r+1}^{N}\sigma^2\mathbf{u}_i\mathbf{u}_i^\dagger \tag{5.4}$$

The inverse covariance matrix is

$$\mathbf{R}_d^{-1} = \sum_{i=1}^{r}(s\lambda_i + \sigma^2)^{-1}\mathbf{u}_i\mathbf{u}_i^\dagger + \sum_{i=r+1}^{N}\sigma^{-2}\mathbf{u}_i\mathbf{u}_i^\dagger \tag{5.5}$$

The previous expression can be rewritten as

$$\begin{aligned}
\mathbf{R}_d^{-1} &= \frac{1}{\sigma^2}\sum_{i=1}^{r}\frac{\mathbf{u}_i\mathbf{u}_i^\dagger}{(1+\frac{s\lambda_i}{\sigma^2})} + \sum_{i=r+1}^{N}\sigma^{-2}\mathbf{u}_i\mathbf{u}_i^\dagger \\
&= \frac{1}{\sigma^2}\sum\nolimits_{i=1}^{N}\mathbf{u}_i\mathbf{u}_i^\dagger - \frac{1}{\sigma^2}\sum_{i=1}^{r}\frac{\frac{s\lambda_i}{\sigma^2}}{(1+\frac{s\lambda_i}{\sigma^2})}\mathbf{u}_i\mathbf{u}_i^\dagger \\
&= \frac{1}{\sigma^2}\Big(\mathbf{I} - \sum_{i=1}^{r}\frac{\frac{s\lambda_i}{\sigma^2}}{(1+\frac{s\lambda_i}{\sigma^2})}\mathbf{u}_i\mathbf{u}_i^\dagger\Big)
\end{aligned} \tag{5.6}$$

For a clutter-to-noise ratio (CNR) much greater than one, i.e. $s\lambda_i \gg \sigma^2$, the inverse of the disturbance covariance matrix can be approximated as [18],

$$\mathbf{R}_d^{-1} \approx \frac{1}{\sigma^2}(\mathbf{I} - \mathbf{P}) \tag{5.7}$$

where

$$\mathbf{P} = \sum\nolimits_{i=1}^{r}\mathbf{u}_i\mathbf{u}_i^\dagger \tag{5.8}$$

is a rank r projection matrix formed with the r eigenvectors corresponding to the r dominant eigenvalues of $\mathbf{R}_d$. Using (5.7) in the NMF test of (3.1), we obtain the LRNMF test

$$\Lambda_{\mathrm{LR}} \equiv \frac{|\mathbf{s}^\dagger(\mathbf{I}-\mathbf{P})\mathbf{x}|^2}{(\mathbf{s}^\dagger(\mathbf{I}-\mathbf{P})\mathbf{s})(\mathbf{x}^\dagger(\mathbf{I}-\mathbf{P})\mathbf{x})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{5.9}$$

It can be noticed that the LRNMF test is invariant to the unknown clutter level s and to the noise power σ^2. Moreover, since $(\mathbf{I}-\mathbf{P})$ is also a projection matrix of rank $(N-r)$, $(\mathbf{I}-\mathbf{P})^2 = (\mathbf{I}-\mathbf{P})$, therefore we can define the transformed vectors $\mathbf{s}_1 = (\mathbf{I}-\mathbf{P})\mathbf{s}$, $\mathbf{x}_1 = (\mathbf{I}-\mathbf{P})\mathbf{x}$ and rewrite the test as

$$\Lambda_{\mathrm{LR}} \equiv \frac{|\mathbf{s}_1^\dagger \mathbf{x}_1|^2}{(\mathbf{s}_1^\dagger \mathbf{s}_1)(\mathbf{x}_1^\dagger \mathbf{x}_1)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{5.10}$$

It is important to observe that the low rank NMF test is the squared cosine of the angle between the transformed steering vector $\mathbf{s}_1$ and the transformed data vector $\mathbf{x}_1$. Fast simulation can be performed using rotation of the primary data vector $\mathbf{x}_1$ as described in Section 3.2.2 for the NMF detector; of course, since the two vectors in this case are not the same as for the NMF case, the IS performance will be different.

5.2.1 FAP approximation: low clutter rank and high CNR

The FAP for the LRNMF detector is derived in [17] and is given by

$$\alpha_{\mathrm{LRNMF}} = (1-\eta)^{N-r-1} \tag{5.11}$$

A crucial point of the LRNMF detector is the clutter rank estimation. The threshold for the LRNMF test, for a fixed FAP, will depend on the rank r of the clutter covariance matrix, specifically it increases with increasing r. A technique for clutter rank estimation can be found in [19]. The case $r = 0$ coincides with the full rank NMF test, which is invariant to the white noise level.

5.3 The LRNMF detector: arbitrary covariance $\mathbf{R}_d$

In this section we derive two alternate expressions for the exact FAP of the LRNMF detector in (5.9). In particular, we do not assume that $(\mathbf{I}-\mathbf{P})$ in (5.7) whitens the primary vector $\mathbf{X}$, as is required for the derivation of the FAP formula in (5.11). These exact forms are valid for *any* primary covariance matrix $\mathbf{R}_d$. The first expression involves singular multivariate Gaussian distributions whereas the second does not.

For ease of notation we denote the projection matrix $(\mathbf{I}-\mathbf{P})$ by $\mathbf{Q}$, which is idempotent. The LRNMF test of (5.9) can then be written as

$$\Lambda_{\mathrm{LR}} \equiv \frac{|\mathbf{s}^\dagger \mathbf{Q}^2 \mathbf{x}|^2}{(\mathbf{s}^\dagger \mathbf{Q}^2 \mathbf{s})(\mathbf{x}^\dagger \mathbf{Q}^2 \mathbf{x})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{5.12}$$

Define the transformed vectors

$$\mathbf{x}_1 = \mathbf{Q}\mathbf{x} \quad \text{and} \quad \mathbf{s}_1 = \mathbf{Q}\mathbf{s} \tag{5.13}$$

Then

$$\mathbf{R}_1 = E\{\mathbf{X}_1\mathbf{X}_1^\dagger\} = \mathbf{Q}E\{\mathbf{X}\mathbf{X}^\dagger\}\mathbf{Q}^\dagger = \mathbf{Q}\mathbf{R}_d\mathbf{Q}^\dagger \tag{5.14}$$

As $\mathbf{Q}$ is a singular matrix, the matrix $\mathbf{R}_1$ is also not of full rank. This follows from the property of the rank of a product of matrices[1]. Hence $\mathbf{X}_1 \sim \mathcal{SCN}_N(0, \mathbf{R}_1)$, where $\mathcal{SCN}$ indicates a singular multivariate Gaussian distribution[2]. Using the above transformations the test becomes

$$\Lambda_{\text{LR}} \equiv \frac{|\mathbf{s}_1^\dagger\mathbf{x}_1|^2}{(\mathbf{s}_1^\dagger\mathbf{s}_1)(\mathbf{x}_1^\dagger\mathbf{x}_1)} = \frac{|\mathbf{s}_1^\dagger\mathbf{x}_1|^2}{\|\mathbf{s}_1\|^2\|\mathbf{x}_1\|^2} = \frac{|(\mathbf{s}_1^\dagger/\|\mathbf{s}_1\|)\ \ \mathbf{x}_1|^2}{\|\mathbf{x}_1\|^2} = \frac{|\mathbf{t}^\dagger\mathbf{x}_1|^2}{\|\mathbf{x}_1\|^2} \tag{5.17}$$

where $\mathbf{t} \equiv \mathbf{s}_1/\|\mathbf{s}_1\|$ is a unit vector.

5.3.1 Exact FAP: using singular Gaussian distributions

Define a unitary transformation $\mathbf{t}_1 = \mathbf{H}\mathbf{t}$, such that the unit vector $\mathbf{t}_1$ has a single element equal to 1 and the remaining $(N-1)$ elements are zero, i.e., $\mathbf{t}_1 = (0, \ldots, 0, 1)^T$. The matrix $\mathbf{H}$ can be an Householder transformation matrix given by

$$\mathbf{H} = \mathbf{I} - \frac{2\mathbf{u}\mathbf{u}^\dagger}{\|\mathbf{u}\|^2} \tag{5.18}$$

and

$$\mathbf{u} = \mathbf{t} + \frac{t_i\|\mathbf{t}\|}{|t_i|}\mathbf{e}_i \tag{5.19}$$

where

$$\mathbf{e}_i = [0, \ldots, 0, 1, 0, \ldots, 0]^T \tag{5.20}$$

[1] $\text{rank}(\mathbf{AB}) \leq \min(\,\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\,)$

[2] A complex Gaussian vector $\mathbf{x}$ of N components with covariance matrix $\mathbf{R}$ and mean vector μ is said to have a singular multivariate Gaussian distribution if $\text{rank}(\mathbf{R}) = p < N$. Using the factorization

$$\mathbf{R} = \mathbf{U}\begin{bmatrix} \mathbf{D}_\lambda & 0 \\ 0 & 0 \end{bmatrix}\mathbf{U}^\dagger \tag{5.15}$$

the covariance matrix can be written as $\mathbf{R} = \mathbf{U}_1\mathbf{D}_\lambda\mathbf{U}_1^\dagger$, where $\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2]$ is an $N \times N$ orthogonal matrix, $\mathbf{U}_1$ is an $N \times p$ column orthogonal matrix, and $\mathbf{D}_\lambda = \text{diag}(\lambda_1, \ldots, \lambda_p)$ with $\lambda_1 > 0$ for $i = 1, \ldots, p$. Define $\mathbf{R}^- \equiv \mathbf{U}_1\mathbf{D}_\lambda^{-1}\mathbf{U}_1^\dagger$, the generalized inverse of $\mathbf{R} = \mathbf{U}_1\mathbf{D}_\lambda\mathbf{U}_1^\dagger$. The p.d.f. of $\mathbf{x}$ is given by

$$f(\mathbf{x}) = \frac{1}{\pi^N|\mathbf{D}_\lambda|}\exp\{(\mathbf{x}-\mu)^\dagger\mathbf{R}^-(\mathbf{x}-\mu)\} \tag{5.16}$$

where $\mathbf{x}$ lies on the p-dimensional linear subspace defined by

$$\mathbf{L}^\dagger(\mathbf{x}-\mu) = 0;\ \mathbf{L} : N \times (N-p),\ \mathbf{L}^\dagger\mathbf{R} = 0,\ \mathbf{L}^\dagger\mathbf{L} = \mathbf{I}_{N-p}$$

has 1 as the ith element and zero elsewhere. The transformation by the matrix $\mathbf{H}$ will eliminate all the elements of $\mathbf{t}$ except the element t_i. Note that $\mathbf{H}$ is an Hermitian unitary matrix

$$\mathbf{H}^{-1} = \mathbf{H}^{\dagger} = \mathbf{H} \tag{5.21}$$

We will construct the $\mathbf{H}$ matrix in order to preserve only the Nth element of the vector $\mathbf{t}$. The likelihood ratio in (5.17) now takes the form

$$\Lambda_{\mathrm{LR}} = \frac{|\mathbf{t}_1^{\dagger}\mathbf{H}\mathbf{x}_1|}{\mathbf{x}_1^{\dagger}\mathbf{x}_1} \tag{5.22}$$

Now define

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{x}_1 \\ \mathbf{R}_y &= E\{\mathbf{y}\mathbf{y}^{\dagger}\} = \mathbf{H}E\{\mathbf{x}_1\mathbf{x}_1^{\dagger}\}\mathbf{H}^{\dagger} = \mathbf{H}\mathbf{R}_1\mathbf{H}^{\dagger} \end{aligned} \tag{5.23}$$

where $\mathbf{y} \sim \mathcal{SCN}_N(0, \mathbf{R}_y)$. The vector $\mathbf{x}_1$ in the test can be replaced by

$$\mathbf{x}_1 = \mathbf{H}^{-1}\mathbf{y} = \mathbf{H}^{\dagger}\mathbf{y} = \mathbf{H}\mathbf{y} \tag{5.24}$$

Then, the likelihood ratio of (5.22) and corresponding test can be rewritten as

$$\Lambda_{\mathrm{LR}} = \frac{|\mathbf{t}_1^{\dagger}\mathbf{y}|}{\mathbf{y}^{\dagger}\mathbf{H}^{\dagger}\mathbf{H}\mathbf{y}} = \frac{|\mathbf{t}_1^{\dagger}\mathbf{y}|}{\mathbf{y}^{\dagger}\mathbf{y}} = \frac{|y_N|^2}{\sum_{i=1}^{N}|y_i|^2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{5.25}$$

where y_i, $i = 1, \ldots, N$ are the elements of the vector $\mathbf{y}$. Now we can write the test in the form

$$\begin{aligned} |y_N|^2 &\underset{H_0}{\overset{H_1}{\gtrless}} \eta\Big(|y_N|^2 + \sum_{i=1}^{N-1}|y_i|^2\Big) \\ |y_N|^2(1-\eta) &\underset{H_0}{\overset{H_1}{\gtrless}} \eta\sum_{i=1}^{N-1}|y_i|^2 \\ |y_N|^2 &\underset{H_0}{\overset{H_1}{\gtrless}} \eta_0\sum_{i=1}^{N-1}|y_i|^2 \\ |y_N| &\underset{H_0}{\overset{H_1}{\gtrless}} \Big(\eta_0\sum\nolimits_{i=1}^{N-1}|y_i|^2\Big)^{1/2} \end{aligned} \tag{5.26}$$

where $\eta_0 = \eta/1 - \eta$.

Define the $(N-1)$ vector $\tilde{\mathbf{y}} \equiv (y_1, \ldots, y_{N-1})^T$, $u_1 \equiv |y_N|$, and

$$b(\eta, \tilde{\mathbf{y}}) \equiv \Big(\eta_0\sum\nolimits_{i=1}^{N-1}|y_i|^2\Big)^{1/2} \tag{5.27}$$

Then, the FAP is given by

$$\begin{aligned}\alpha_{\text{LRNMF}} &= P\Big(|Y_N| \geq \Big(\eta_0 \sum\nolimits_{i=1}^{N-1} |Y_i|^2\Big)^{1/2}\Big) \\ &= E\Big\{P\Big(U_1 \geq b(\eta, \tilde{\mathbf{Y}})\big|\tilde{\mathbf{Y}}\Big)\Big\} \\ &\triangleq E\big\{g(\eta, \tilde{\mathbf{Y}})\big\} \end{aligned} \tag{5.28}$$

and the g function can be expressed in integral form as

$$g(\eta, \tilde{\mathbf{y}}) = \int_{b(\eta,\tilde{\mathbf{y}})}^{\infty} f(u_1|\tilde{\mathbf{y}}) du_1 \tag{5.29}$$

The conditional distribution of Y_N is also Gaussian and we have to determine the conditional mean and variance. The vector $\mathbf{y}$ can be written as $\mathbf{y} = [\tilde{\mathbf{y}}; y_N]$ and the covariance matrix of $\mathbf{y}$ can be partitioned as

$$\mathbf{R}_y = \begin{bmatrix} E\{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^\dagger\} & E\{\tilde{\mathbf{Y}}Y_N^\dagger\} \\ E\{Y_N\tilde{\mathbf{Y}}^\dagger\} & E\{Y_N Y_N^\dagger\} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{\tilde{\mathbf{y}}} & \mathbf{R}_{\tilde{\mathbf{y}}y_N} \\ \mathbf{R}_{y_N\tilde{\mathbf{y}}} & \sigma^2_{y_N} \end{bmatrix} \tag{5.30}$$

where $\mathbf{R}_{\tilde{\mathbf{y}}}$ is the $(N-1)\times(N-1)$ singular covariance matrix of the vector $\tilde{\mathbf{y}}$ and $\sigma^2_{y_N}$ is the variance of the random variable Y_N. The conditional mean and variance of Y_N are

$$\begin{aligned} E\{Y_N|\tilde{\mathbf{y}}\} &= \mathbf{R}_{y_N\tilde{\mathbf{y}}}\mathbf{R}^-_{\tilde{\mathbf{y}}}\tilde{\mathbf{y}} = \mathbf{c}^T\tilde{\mathbf{y}} = \sum_{i=1}^{N-1} c_i\, y_i \\ \text{var}\{Y_N|\tilde{\mathbf{y}}\} &= \sigma^2_{y_N} - \mathbf{R}_{y_N\tilde{\mathbf{y}}}\mathbf{R}^-_{\tilde{\mathbf{y}}}\mathbf{R}_{\tilde{\mathbf{y}}y_N} \end{aligned} \tag{5.31}$$

where $\mathbf{c}^T \equiv \mathbf{R}_{y_N\tilde{\mathbf{y}}}\mathbf{R}^-_{\tilde{\mathbf{y}}}$, $c_i, i = 1, \ldots, N-1$ are the elements of the vector $\mathbf{c}$, and $\mathbf{R}^-_{\tilde{\mathbf{y}}}$ is the g-inverse of $\mathbf{R}_{\tilde{\mathbf{y}}}$.

Conditioned on the vector $\tilde{\mathbf{Y}}$, the random variable Y_N is Gaussian with mean and variance given in the expressions above. Then the random variable $U_1 = |Y_N|$ has a Rice distribution with noncentrality parameter

$$s^2 = \Big|\sum\nolimits_{i=1}^{N-1} c_i\, y_i\Big|^2 \tag{5.32}$$

and parameter

$$\sigma^2 = \text{var}\{Y_N|\tilde{\mathbf{y}}\}/2 \tag{5.33}$$

Hence the integral in (5.29) can be written as

$$g(\eta, \tilde{\mathbf{y}}) = \int_{b(\eta,\tilde{\mathbf{y}})}^{\infty} \frac{u_1}{\sigma^2} \exp\big[-(u_1^2 + s^2)/2\sigma^2\big] I_0(u_1 s/\sigma^2)\, du_1 \tag{5.34}$$

and the FAP becomes

$$\boxed{\alpha_{\text{LRNMF}} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{b(\eta,\tilde{\mathbf{y}})}^{\infty} \frac{u_1}{\sigma^2} \exp\left[-(u_1^2+s^2)/2\sigma^2\right] I_0(u_1 s/\sigma^2)\, f(\tilde{\mathbf{y}})\, du_1 d\tilde{\mathbf{y}}} \tag{5.35}$$

where $f(\tilde{\mathbf{y}})$ is the density corresponding to the singular Gaussian distribution $\mathcal{SCN}_{N-1}(0, \mathbf{R}_{\tilde{\mathbf{y}}})$. This is one expression for the FAP of the LRNMF detector for *arbitrary covariance matrix* $\mathbf{R}_d$ of the primary data vector.

5.3.2 Exact FAP: using nonsingular Gaussian distributions

We start with the covariance matrix $\mathbf{R}_1$ of $\mathbf{X}_1 \sim \mathcal{SCN}_N(0, \mathbf{R}_1)$ having some rank say $p < N$ in (5.14) and write its decomposition as

$$\mathbf{R}_1 = \mathbf{Q}\mathbf{R}_d\mathbf{Q}^\dagger = \mathbf{V}\Lambda\mathbf{V}^\dagger \tag{5.36}$$

where $\mathbf{V}$ is the matrix of eigenvectors corresponding to the eigenvalues $\{\lambda_i\}_1^N$ in the diagonal matrix Λ, which can be written as

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & . & . & . & 0 \\ 0 & . & . & . & . & . \\ . & . & \lambda_p & . & . & . \\ . & . & . & 0 & . & . \\ . & . & . & . & . & 0 \\ 0 & . & . & . & 0 & 0 \end{pmatrix} \tag{5.37}$$

Define

$$\Lambda_p \equiv \begin{pmatrix} \lambda_1 & 0 & . & . & 0 \\ 0 & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & 0 \\ 0 & . & . & 0 & \lambda_p \end{pmatrix} \tag{5.38}$$

and

$$\begin{aligned} \mathbf{V}_p &\equiv \text{upper left block of } \mathbf{V} \in R^{p\times p} \\ \mathbf{V}_{N-p} &\equiv \text{lower left block of } \mathbf{V} \in R^{(N-p)\times p} \end{aligned} \tag{5.39}$$

Using these definitions we can write $\mathbf{R}_1$ as

$$\mathbf{R}_1 = \begin{pmatrix} \mathbf{V}_p\,\Lambda_p\,\mathbf{V}_p^\dagger & \mathbf{V}_p\,\Lambda_p\,\mathbf{V}_{N-p}^\dagger \\ \mathbf{V}_{N-p}\,\Lambda_p\,\mathbf{V}_p^\dagger & \mathbf{V}_{N-p}\,\Lambda_p\,\mathbf{V}_{N-p}^\dagger \end{pmatrix} \tag{5.40}$$

which also turns out to be the covariance matrix of a new vector $\mathbf{X}_1$ defined as

$$\mathbf{X}_1 \equiv \underbrace{\begin{bmatrix} \mathbf{V}_p \\ \mathbf{V}_{N-p} \end{bmatrix}}_{N\times p} \Lambda_p^{1/2}\,\mathbf{W}; \quad \mathbf{W} \sim \mathcal{CN}_p(0, \mathbf{I}) \tag{5.41}$$

and hence this new $\mathbf{X}_1$ is statistically identical to the vector $\mathbf{X}_1$ in the beginning of this subsection. This is a well known representation ([20]) for expressing a singular Gaussian vector in terms of a nonsingular Gaussian vector with independent components.

Substituting this definition into (5.17) yields the likelihood ratio

$$\begin{aligned}
\Lambda_{\mathrm{LR}} &= \frac{|\mathbf{t}^\dagger \mathbf{x}_1|^2}{\|\mathbf{x}_1\|^2} \\
&= \frac{\left| \mathbf{t}^\dagger \begin{bmatrix} \mathbf{V}_p \\ \mathbf{V}_{N-p} \end{bmatrix} \Lambda_p^{1/2}\, \mathbf{W} \right|^2}{\mathbf{W}^\dagger \Lambda_p^{1/2} \begin{bmatrix} \mathbf{V}_p^\dagger & \mathbf{V}_{N-p}^\dagger \end{bmatrix} \begin{bmatrix} \mathbf{V}_p \\ \mathbf{V}_{N-p} \end{bmatrix} \Lambda_p^{1/2}\, \mathbf{W}} \\
&= \frac{\left| \mathbf{t}^\dagger \begin{bmatrix} \mathbf{V}_p \\ \mathbf{V}_{N-p} \end{bmatrix} \Lambda_p^{1/2}\, \mathbf{W} \right|^2}{\mathbf{W}^\dagger \Lambda_p \mathbf{W}}
\end{aligned} \tag{5.42}$$

Now define

$$\mathbf{Y} \equiv \Lambda_p^{1/2} \mathbf{W} \tag{5.43}$$

and the $p \times 1$ unit vector

$$\mathbf{t}_1 \equiv \begin{bmatrix} \mathbf{V}_p^\dagger & \mathbf{V}_{N-p}^\dagger \end{bmatrix} \mathbf{t} \tag{5.44}$$

Then $Y \sim \mathcal{CN}_p(0, \mathbf{\Lambda_p})$ and the likelihood ratio becomes

$$\Lambda_{\mathrm{LR}} = \frac{|\mathbf{t}_1^\dagger \mathbf{y}|^2}{\mathbf{y}^\dagger \mathbf{y}} \tag{5.45}$$

Next we find a unitary transformation $\mathbf{H}$ (not to be confused with the unitary transformation $\mathbf{H}$ defined on page 78) such that

$$\mathbf{e}_1 \equiv \mathbf{H}\, \mathbf{t}_1 = \begin{pmatrix} 1 & \cdot & \cdot & \cdot & 0 \end{pmatrix}^T \tag{5.46}$$

Then the likelihood ratio can be written as

$$\Lambda_{\mathrm{LR}} = \frac{|\mathbf{e}_1^\dagger \mathbf{H} \mathbf{y}|^2}{\mathbf{y}^\dagger \mathbf{y}} \tag{5.47}$$

Defining

$$\mathbf{Z} \equiv \mathbf{H}\, \mathbf{Y} \tag{5.48}$$

leads to

$$\Lambda_{\mathrm{LR}} = \frac{|\mathbf{e}_1^\dagger \mathbf{z}|^2}{\mathbf{z}^\dagger \mathbf{z}} \tag{5.49}$$

with $Z \sim \mathcal{CN}_p(0, \mathbf{R}_z)$ where

$$\mathbf{R}_z = \mathbf{H} \Lambda_p \mathbf{H}^\dagger \tag{5.50}$$

The likelihood ratio test can therefore be written as

$$\frac{|z_1|^2}{\sum_{i=1}^{p} |z_i|^2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{5.51}$$

where the z_i's denote the components of $\mathbf{z}$. Rearranging yields

$$|z_1| \underset{H_0}{\overset{H_1}{\gtrless}} \left(\eta_o \sum\nolimits_{i=2}^{p} |z_i|^2\right)^{1/2} \tag{5.52}$$

where $\eta_o = \eta/(1-\eta)$, which can be considered (somewhat loosely speaking) as the nonsingular version of the test in (5.26) on page 79. The rest of the derivation for FAP is similar to those in Section 5.3.1 but we go through the steps.

Defining $\tilde{\mathbf{z}} \equiv \begin{pmatrix} z_2 & \cdot & \cdot & \cdot & z_p \end{pmatrix}^T$, the covariance matrix $\mathbf{R}_z$ can be partitioned as

$$\begin{aligned}
\mathbf{R}_z &= E\left\{ \begin{bmatrix} Z_1 \\ \tilde{\mathbf{Z}} \end{bmatrix} \begin{bmatrix} Z_1^\star & \tilde{\mathbf{Z}}^\dagger \end{bmatrix} \right\} \\
&= \begin{bmatrix} E\{|Z_1|^2\} & E\{Z_1 \tilde{\mathbf{Z}}^\dagger\} \\ E\{\tilde{\mathbf{Z}} Z_1^\star\} & E\{\tilde{\mathbf{Z}}\tilde{\mathbf{Z}}^\dagger\} \end{bmatrix} \\
&\triangleq \begin{bmatrix} \sigma_1^2 & \mathbf{R}_{z_1\tilde{\mathbf{z}}} \\ \mathbf{R}_{\tilde{\mathbf{z}}z_1} & \mathbf{R}_{\tilde{\mathbf{z}}} \end{bmatrix}
\end{aligned} \tag{5.53}$$

The random variable Z_1 is conditionally Gaussian with density $f(z_1 \mid \tilde{\mathbf{z}})$ corresponding to the distribution $\mathcal{CN}_1(\mu_c, \sigma_c^2)$ where the conditional mean and variance μ_c and σ_c^2 are given by

$$\begin{aligned}
\mu_c = E\{Z_1 \mid \tilde{\mathbf{z}}\} &= \mathbf{R}_{z_1\tilde{\mathbf{z}}} \mathbf{R}_{\tilde{\mathbf{z}}}^{-1} \tilde{\mathbf{z}} \\
\sigma_c^2 = \mathrm{var}(Z_1 \mid \tilde{\mathbf{z}}) &= \sigma_1^2 - \mathbf{R}_{z_1\tilde{\mathbf{z}}} \mathbf{R}_{\tilde{\mathbf{z}}}^{-1} \mathbf{R}_{\tilde{\mathbf{z}}z_1}
\end{aligned} \tag{5.54}$$

The random variable $U \equiv |Z_1|$ is then conditionally Rice with density

$$f(u \mid \tilde{\mathbf{z}}) = \frac{u}{\sigma^2} \exp\big[-(u^2+s^2)/2\sigma^2\big] I_0(us/\sigma^2), \quad u \geq 0 \tag{5.55}$$

where $s = |\mu_c|$ and $\sigma^2 = \sigma_c^2/2$. Defining

$$b(\eta, \tilde{\mathbf{z}}) \equiv \left(\eta_0 \sum\nolimits_{i=2}^{p} |z_i|^2\right)^{1/2} \tag{5.56}$$

the FAP of the test in (5.52) can be expressed as

$$\begin{aligned}
\alpha_{\mathrm{LRNMF}} &= P\Big(|Z_1| \geq \left(\eta_0 \sum\nolimits_{i=2}^{p} |Z_i|^2\right)^{1/2}\Big) \\
&= E\Big\{P\Big(U \geq b(\eta, \tilde{\mathbf{Z}}) \big| \tilde{\mathbf{Z}}\Big)\Big\}
\end{aligned}$$

$$\triangleq \quad E\{g(\eta, \tilde{\mathbf{Z}})\} \tag{5.57}$$

where

$$g(\eta, \tilde{\mathbf{z}}) = \int_{b(\eta,\tilde{\mathbf{z}})}^{\infty} \frac{u}{\sigma^2} \exp\big[-(u^2+s^2)/2\sigma^2\big] I_0(us/\sigma^2)\, du \tag{5.58}$$

using (5.55). Then, similarly to (5.35), the FAP can be written as

$$\boxed{\alpha_{\text{LRNMF}} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{b(\eta,\tilde{\mathbf{z}})}^{\infty} \frac{u}{\sigma^2} \exp\big[-(u^2+s^2)/2\sigma^2\big] I_0(us/\sigma^2)\, f(\tilde{\mathbf{z}})\, du\, d\tilde{\mathbf{z}}} \tag{5.59}$$

where $f(\tilde{\mathbf{z}})$ corresponds to the (nonsingular) density of $\mathcal{CN}_{p-1}(0, \mathbf{R}_{\tilde{\mathbf{z}}})$. This is the second expression for the exact FAP of the LRNMF detector for *arbitrary covariance matrix* $\mathbf{R}_d$ of the primary data vector.

5.4 Nominal statistical model for simulation and threshold setting

A nominal statistical model for the radar returns in the (target-free) primary data vector is required in order to be able to specify a covariance matrix $\mathbf{R}_d$ under the Gaussian assumption. Such a model can then be used to derive threshold settings for desired FAP values for the detector either through simulation or through direct computation of FAP expressions.

For the disturbance covariance $\mathbf{R}_d = s\mathbf{R}_c + \sigma^2\mathbf{I}$ in (5.2) we start again with the low clutter-rank approximation of (5.4) on page 76 as

$$\begin{aligned} \mathbf{R}_d &= \sum_{i=1}^{r} (s\lambda_i + \sigma^2)\, \mathbf{u}_i \mathbf{u}_i^\dagger + \sum_{i=r+1}^{N} \sigma^2\, \mathbf{u}_i \mathbf{u}_i^\dagger \\ &= \mathbf{U}\mathbf{D}\mathbf{U}^\dagger \end{aligned} \tag{5.60}$$

and identify the matrix of eigenvalues $\mathbf{D}$ as

$$\begin{aligned} \mathbf{D} &= \text{diag}\,\big(\, s\lambda_1 + \sigma^2, \ldots, s\lambda_r + \sigma^2, \sigma^2, \ldots, \sigma^2 \,\big) \\ &= \sigma^2\, \text{diag}\,\big(\, 1 + s\lambda_1/\sigma^2, \ldots, 1 + s\lambda_r/\sigma^2, \underbrace{1, \ldots, 1}_{N-r} \,\big) \end{aligned} \tag{5.61}$$

Now, choice of the unitary matrix $\mathbf{U}$ of eigenvectors will determine the covariance matrix $\mathbf{R}_d$. The simplest choice that can be made (at the risk of seeming somewhat subjective) is $\mathbf{U} = \mathbf{I}$. In such a case, $\mathbf{R}_d = \mathbf{D}$. That is

$$\mathbf{R}_d = \mathbf{D} = \sigma^2 \begin{pmatrix} 1 + s\lambda_1/\sigma^2 & 0 & . & . & . & 0 \\ 0 & . & . & . & . & . \\ . & . & 1 + s\lambda_r/\sigma^2 & . & . & . \\ . & . & . & 1 & . & . \\ . & . & . & . & . & 0 \\ 0 & . & . & . & 0 & 1 \end{pmatrix} \tag{5.62}$$

It turns out that this choice is not without some practical significance. In a special case, using the high clutter-to-noise ratio assumption

$$\frac{s\lambda_i}{\sigma^2} \gg 1$$

with

$$\sigma^2 = 1;\ \lambda_i = 1,\ i = 1, \ldots, r$$

leads to a covariance matrix of the form

$$\mathbf{R}_d = \mathbf{D} = \begin{pmatrix} s & 0 & . & . & 0 \\ 0 & s & . & . & . \\ . & . & . & . & . \\ . & . & . & 1 & 0 \\ 0 & . & . & 0 & 1 \end{pmatrix} \tag{5.63}$$

A plot of the eigenspectrum of $\widehat{\mathbf{R}}_d$ estimated from Gaussian data generated according to the above $\mathbf{R}_d$ is shown in Figure 5.1. This eigenspectrum shows a close resemblance to the eigenspectrum estimated from KASSPER data, [17]. We note of course that this property of our model is due to the shape of the matrix of eigenvalues $\mathbf{D}$ (rather than the particular choice $\mathbf{U} = \mathbf{I}$).

A consequence of the structure of the nominal covariance matrix $\mathbf{R}_d$ in (5.62) is described in the following subsection.

5.4.1 Exact FAP of LRNMF detector: nominal $\mathbf{R}_d$

We calculate here the FAP of the LRNMF detector when the covariance matrix $\mathbf{R}_d$ of the primary data vector is as in (5.62). The assumption of high clutter-to-noise ratio is not used. Therefore, the only restriction is on the shape of $\mathbf{R}_d$. As $\mathbf{R}_d$ is diagonal, it follows that its eigenvector matrix $\mathbf{U}$ are composed of the columns of $\mathbf{I}$ and hence, for a given r, the projection matrix $\mathbf{Q}$ is given by

$$\begin{aligned} \mathbf{Q} &= \mathbf{I} - \sum_{i=1}^{r} u_i u_i^\dagger \\ &= \sum_{i=1}^{N-r} u_i u_i^\dagger \\ &= \begin{pmatrix} 0 & . & . & . & . & 0 \\ . & . & . & . & . & . \\ . & . & 0 & . & . & . \\ . & . & . & 1 & . & . \\ . & . & . & . & . & 0 \\ 0 & . & . & . & 0 & 1 \end{pmatrix} \end{aligned} \tag{5.64}$$

which is of rank $N - r$. Applying this and $\mathbf{R}_d$ in (5.62) to (5.36) yields

$$\mathbf{R}_1 = \mathbf{V}\Lambda\mathbf{V}^\dagger$$

$$= \mathbf{Q} \tag{5.65}$$

This implies, from (5.37), that

$$\Lambda = \begin{pmatrix} 1 & 0 & . & . & . & 0 \\ 0 & . & . & . & . & . \\ . & . & 1 & . & . & . \\ . & . & . & 0 & . & . \\ . & . & . & . & . & 0 \\ 0 & . & . & . & 0 & 0 \end{pmatrix} \tag{5.66}$$

$p = N - r$, and thus Λ_p in (5.38) is $\Lambda_p = \mathbf{I}_{N-r}$. Using this in (5.50) yields $\mathbf{R}_z = \mathbf{I}_{N-r}$. Therefore the test in (5.52) represents a simple CA-CFAR test with FAP given by

$$\alpha_{\text{LRNMF}} = (1-\eta)^{N-r-1}$$

which coincides with the expression for the LRNMF detector in (5.11) on page 77.

Therefore we can conclude that whereas the above formula is an approximation (albeit a good one) for the FAP under the assumptions of low clutter-rank and high clutter-to-noise ratios, it is an exactitude if the covariance matrix of the data possesses the structure in (5.62). In the latter case the clutter-to-noise ratio does not matter.

5.5 The LRNAMF detector

The low-rank normalized adaptive filter (LRNAMF) detector is described by

$$\Lambda_{\text{LR-A}} \equiv \frac{|\mathbf{s}^\dagger(\mathbf{I}-\widehat{\mathbf{P}})\mathbf{x}|^2}{(\mathbf{s}^\dagger(\mathbf{I}-\widehat{\mathbf{P}})\mathbf{s})(\mathbf{x}^\dagger(\mathbf{I}-\widehat{\mathbf{P}})\mathbf{x})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{5.67}$$

where

$$\widehat{\mathbf{P}} = \sum\nolimits_{i=1}^{r} \mathbf{u}_i \mathbf{u}_i^\dagger \tag{5.68}$$

and $\{\mathbf{u}_i\}_1^N$ are eigenvectors of the covariance matrix estimate $\widehat{\mathbf{R}}_d$ obtained from the secondary vectors with r being the estimated rank of the clutter component of the disturbance.

5.5.1 FAP approximation: low clutter rank and high CNR

The main difference between the LRNMF and LRNAMF detectors (the former given in (5.9)), for a fixed value of r used in both detectors, lies in the set of eigenvectors $\{\mathbf{u}_i\}_1^N$ used to compose the matrix $\mathbf{P}$ (or $\widehat{\mathbf{P}}$). However, for the low-rank clutter model of (5.4) and subject to the high clutter-to-noise ratio approximation of (5.7), the FAP performance of the LRNMF detector is given by the expression of (5.11) and this is *independent of the eigenvectors* used in the detector. This implies that the LRNMF detector is CFAR with respect to the structure of covariance matrices $\mathbf{R}_d$ of data that arise from models satisfying

these assumptions. Therefore, for data having the same eigenspectrum as in the above model, and for the same value of r used in both detectors, the FAP performance of the LRNAMF detector will be, largely, equal to that of the LRNMF detector and given by the same formula for each realization of sample covariance matrix $\widehat{\mathbf{R}}_d$. The LRNAMF detector will also be approximately CFAR under the above assumptions on clutter rank and power. Nevertheless, it must be emphasized that the FAP expression of (5.11) is an approximation.

The above conclusion is borne out by simulation results.

5.5.2 Exact FAP: arbitrary covariance $\mathbf{R}_d$

If the model assumptions do not hold, specifically in cases where the CNR is not high, then the expressions for exact FAP of the LRNMF detector derived in (5.35) and (5.59) can be used to determine the exact FAP of the LRNAMF detector. Whereas the influence of the secondary vectors on the LRNAMF detector is felt through the eigenvectors in $\widehat{\mathbf{P}}$, via singular value decomposition of the covariance matrix estimate $\widehat{\mathbf{R}}_d$, the FAP is affected by the covariance matrix $\mathbf{R}_{\mathbf{z}}$ in the expression (5.59). This matrix is derived from the eigenvectors in $\widehat{\mathbf{P}}$ through a series of transformation described in Sections 5.3 and 5.3.2 and is, strictly speaking, a random matrix whose statistical properties are related to those of the secondary vectors. Therefore the FAP expression for the LRNMF detector in (5.59) can be considered as a conditional probability expression; the average of this over the statistics of the covariance matrix $\mathbf{R}_{\mathbf{z}}$ will produce the exact FAP performance of the LRNAMF detector. Although it does not appear to be analytically tractable, this fact can be formally expressed by

$$\begin{aligned} \alpha_{\text{LRNAMF}} &= E_{\mathbf{R}_{\mathbf{z}}}\{\alpha_g(\mathbf{R}_{\mathbf{z}})\} \\ &= E_{\mathbf{R}_{\mathbf{z}}}\{E\{g(\eta, \tilde{\mathbf{Z}})\big|\mathbf{R}_{\mathbf{z}}\}\} \end{aligned} \tag{5.69}$$

with $g(\eta, \tilde{\mathbf{z}})$ given in (5.58).

5.6 Simulations for LR detectors

We describe here the simulation results obtained for FAP and threshold estimation for both LRNMF and LRNAMF detectors for various assumptions on the data covariance matrix $\mathbf{R}_d$. The two FAP expressions, (5.35) and (5.59), have been used to set up estimators for FAP. The estimator for the LRNMF detector can be written as

$$\widehat{\alpha}_{\text{LRNMF}} = \frac{1}{K}\sum_{1}^{K} g(\eta, \tilde{\mathbf{Z}}); \quad \tilde{\mathbf{Z}} \sim \mathcal{CN}_{p-1}(0, \mathbf{R}_{\tilde{\mathbf{z}}}) \tag{5.70}$$

For simplicity we have not used any IS. The FAP estimator for the LRNAMF detector is configured as

$$\begin{aligned}\widehat{\alpha}_{\text{LRNAMF}} &= \frac{1}{K_2}\sum_{1}^{K_2}\Big[\frac{1}{K_1}\sum_{1}^{K_1}[g(\eta,\tilde{\mathbf{Z}})]^{(i)}\Big]^{(j)};\\ &\qquad \tilde{\mathbf{Z}}\sim\mathcal{CN}_{p-1}(0,\mathbf{R}_{\tilde{\mathbf{z}}}),\ \mathbf{R}_{\tilde{\mathbf{z}}}\sim f(\text{secondary vectors}) \quad (5.71)\end{aligned}$$

where the inner simulation is conditioned on the covariance matrix $\mathbf{R}_{\tilde{\mathbf{z}}}$, which depends on the data covariance $\mathbf{R}_d$ in place.

Detector thresholds have been estimated using the inverse method.

5.6.1 $\mathbf{R}_d$ models used

Three different data covariance models have been used to study the performances of LRNMF and LRNAMF detectors for FAP as well as detection probability. The models are

Case 1 : In this model, the disturbance is assumed to be white Gaussian noise which ideally has a flat eigenspectrum.

Case 2 In this model we use an eigenspectrum with eigenvalues ranging from 10^4 to 0.1, covering the same range as Model 1, but linearly (in dB) with eigenvalue number.

Case 3 This is the nominal disturbance model which has the matrix of eigenvalues described in (5.63), and corresponds to the case of low clutter-rank and high CNR.

Data generated using these models have estimated eigenspectra shown in Figure 5.1. Wherever detection probability results are provided for a certain model, the detector thresholds used have been estimated to provide the stated FAPs for that model.

5.6.2 Results for FAP

The first result for FAP of the LRNMF detector is shown in Figure 5.2 for the eigenspectrum of Case 3. As expected, the simulation results using the g-method coincide with the formula of (5.11), indicated by stars in the figure. Corresponding simulation gains for the g-method are shown in Figure 5.3.

The next figure, Figure 5.4 shows results for FAP versus thresholds of the LRNMF and LRNAMF detectors for the three eigenspectra models. Lack of CFAR is clearly observed for both the LRNMF and LRNAMF detectors with change in the covariance structure of the data. However, a surprising observation is that for the same data eigenspectrum, the *LRNMF and LRNAMF*

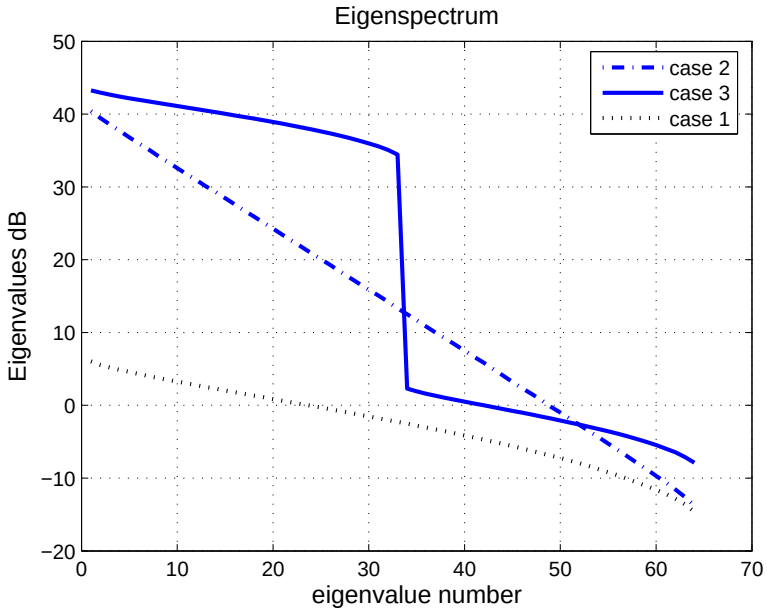

Figure 5.1: Estimated eigenspectrum for the 3 data models. Clutter rank is $r = 33$.

detectors match very closely in terms of achieved FAP values for given thresholds[3]. A plausible explanation for this behaviour lies in the following arguments. Firstly, LR detectors seem to ignore some of the information available in the secondary vectors, depending solely on the eigenvectors obtained through spectral decomposition and not on the actual estimated eigenvalues. Furthermore, for a sufficiently large number L of secondary vectors, the estimated eigenspectrum resembles the actual eigenspectrum in shape. This fact may manifest itself through the eigenvectors of the spectral decomposition. We emphasize that although the match between LRNMF and LRNAMF detectors appears close graphically, numerical estimates obtained for the two detectors do differ.

Estimated simulation gains for the LRNAMF detector are shown in Figure 5.5.

5.6.3 Results for detection probability

Various detection probability results are shown in Figures 5.6 - 5.9. Figure 5.6 shows that the detection probability of the LRNAMF detector is almost unchanged from the homogeneous case in the presence of interfering targets in the

[3]The exception appears to be Case 2. However this is not a departure from the above observation as the number of simulation trials for FAPs below 2×10^{-4} were insufficient, according to gain estimates that were made. Hence more simulations are required to make a definitive conclusion for this eigenspectrum.

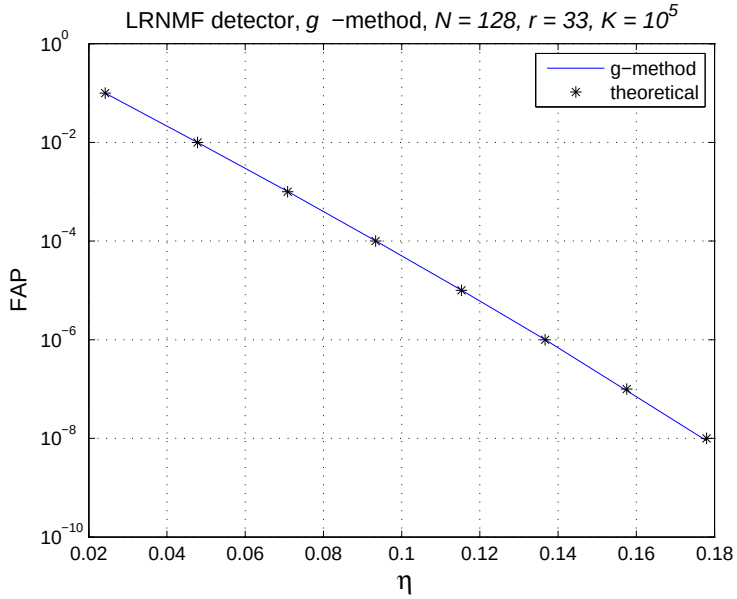

Figure 5.2: Comparison of the g-method estimator and formula (5.11) for FAP. $N = 128, r = 33$. Eigenspectrum is for Case 3.

secondary data, for both Swerling 0 and 1 target models. Comparisons with other detectors are shown in Figure 5.7 and Figure 5.8 for the eigenspectrum of Case 3. Comparing the performance of the adaptive detectors (LRNAMF and NAMF) for this model, the LR version performs better than the NAMF detector for both Swerling 0 and Swerling 1 target models, not only in the homogeneous case but also in the presence of interfering targets contaminating the secondary vectors. On the other hand when the background is as in Case 1 (i.e., clutter has rank zero), the NAMF detector performs better than the LRNAMF detector in the presence of interfering targets, as shown in Figure 5.9 and Figure 5.10. In homogeneous background the NAMF detector performs better for the Swerling 0 target model and the two detectors perform the same for Swerling 1 target model.

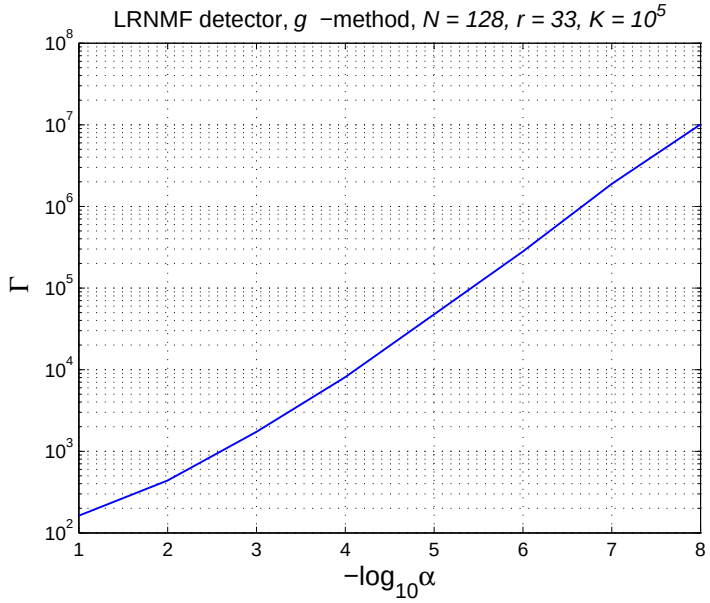

Figure 5.3: Gain of the g-method for $N = 128, r = 33$. Eigenspectrum is for Case 3.

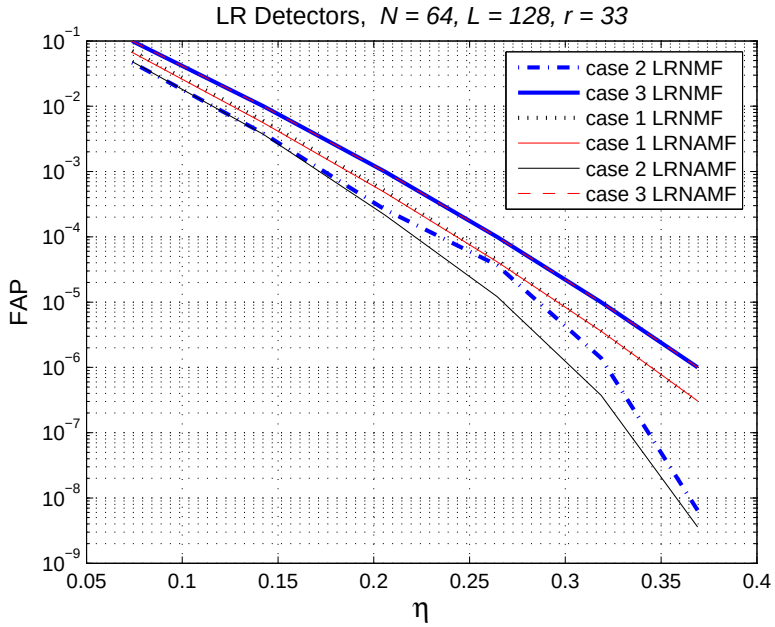

Figure 5.4: Thresholds versus FAP for three eigenspectrum models.

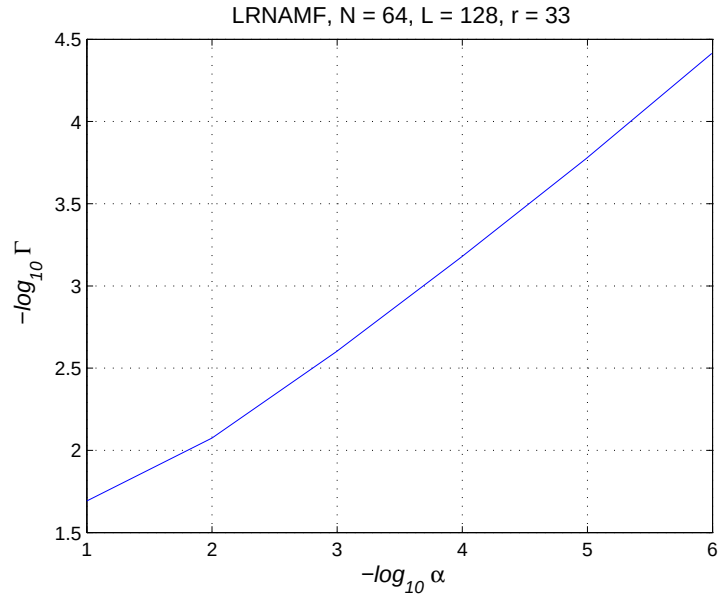

Figure 5.5: Gain of the g-method for $N = 64, L = 128, r = 33$. Eigenspectrum is for Case 3.

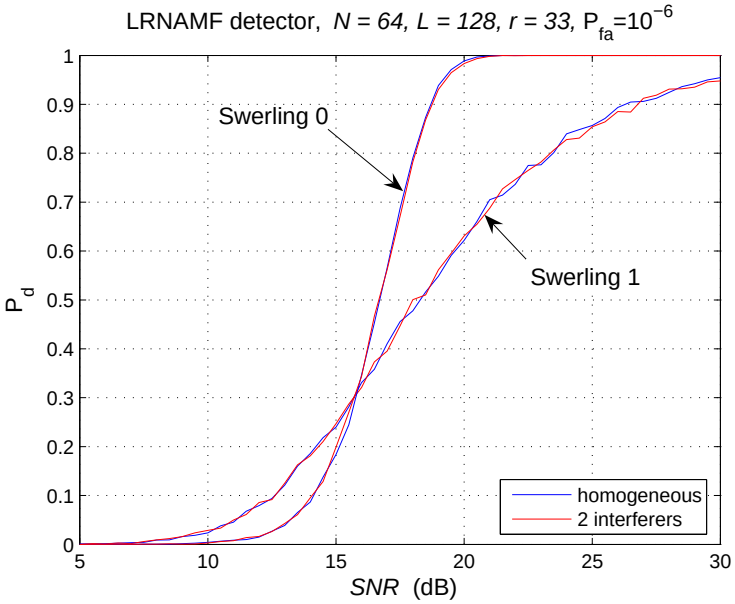

Figure 5.6: Detection probability for LRNAMF detector, $N = 64, L = 128$. Eigenspectrum is for Case 3.

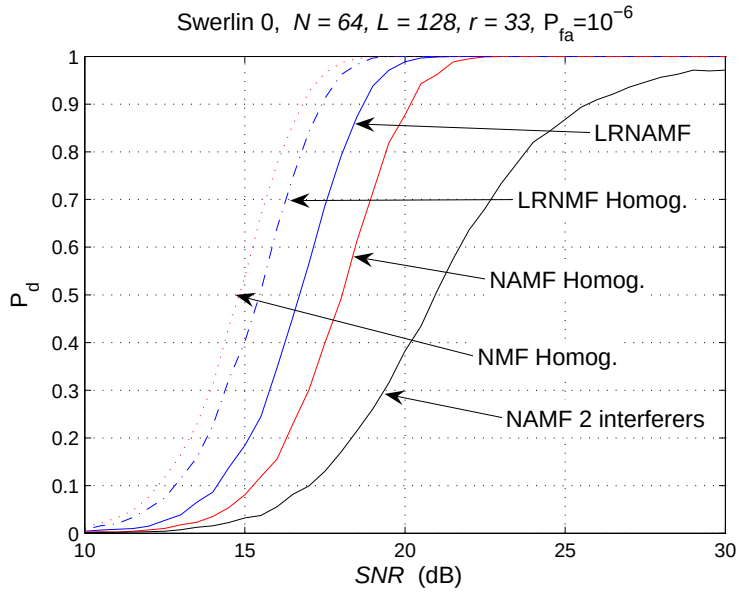

Figure 5.7: Detection probability comparisons, $N = 64, L = 128$. Eigenspectrum is for Case 3.

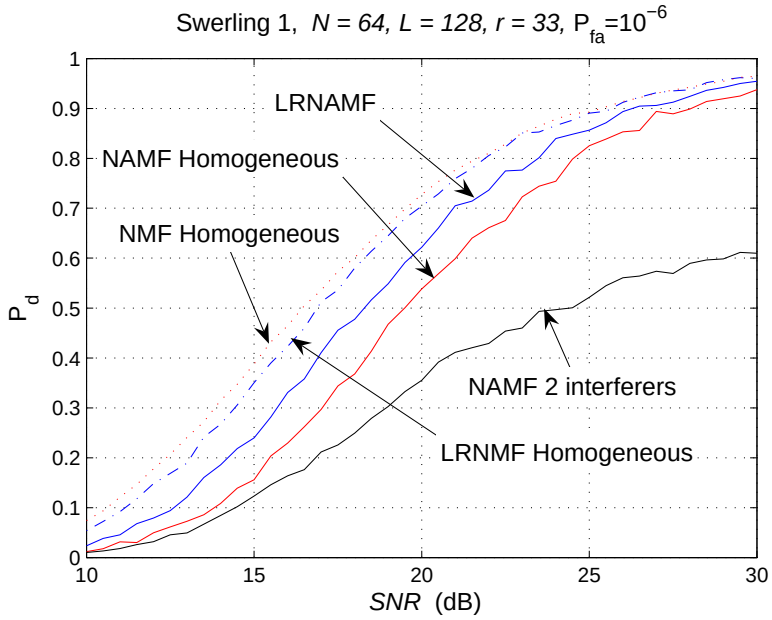

Figure 5.8: Detection probability comparisons, $N = 64, L = 128$. Eigenspectrum is for Case 3.

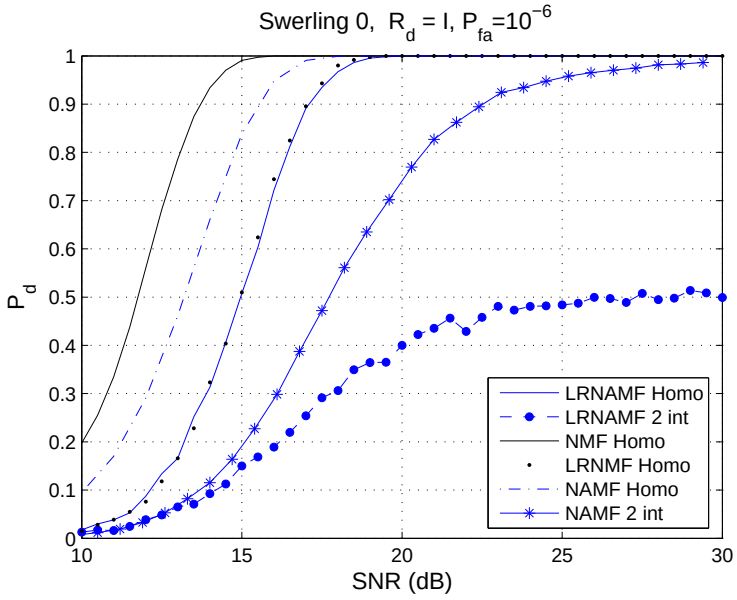

Figure 5.9: Detection probability comparisons (Swerling 0), $N = 64, L = 128$. Eigenspectrum is for Case 1.

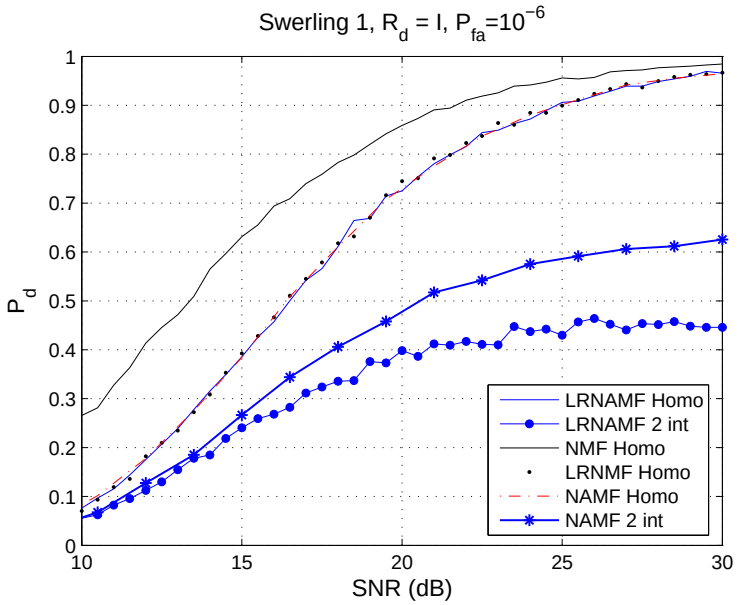

Figure 5.10: Detection probability comparisons (Swerling 1), $N = 64, L = 128$. Eigenspectrum is for Case 1.

5.7 Conclusion

In this chapter some theoretical investigations have been made into the performance of low-rank STAP detectors. Principally, we have been able to characterize the FAP performance of LRNAMF detectors in terms of detection thresholds and disturbance backgrounds. Several detection probability simulation results have also been obtained. Some of these results are surprising and might seem somewhat counterintuitive. We have not been able to find complete justifications for these observations and clearly further work is required.

Bibliography

[1] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerospace and Electronic Systems*, AES-28, NO. 1, pp. 208-216, January 1992.

[2] E.J. Kelly, "An adaptive detection algorithm", *IEEE Trans. Aerospace and Electronic Systems*, AES-22, NO. 1, pp. 115-, March 1986.

[3] M. Rangaswamy, J. H. Michels, and B. Himed, "Statistical analysis of the non-homogeneity detector for STAP applications," *Digital Signal Processing*, vol. 14, pp. 253-267, 2004.

[4] D. Maiwald and D. Kraus, "Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices", *IEE Proc. Radar, Sonar Navig.*, Vol. 147, NO.4, August 2000.

[5] R. Srinivasan, *Importance Sampling - Applications in Communications and Detection.* Berlin: Springer-Verlag, 2002.

[6] R. Srinivasan, "Some results in importance sampling and an application to detection," *Signal Processing*, vol 65, Issue 1, pp 73-88, February 1998.

[7] S. R. Babu and R. Srinivasan, "Analysis of envelope detected mean level CFAR processors using importance sampling", *The record of the IEEE Radar Conference*, May 2000.

[8] P. Billingsley
Convergence of Probability Measures, John Wiley, 1999.

*[9] R. Srinivasan and M. Rangaswamy, "Fast estimation of false alarm probabilities of STAP detectors - the AMF", Proceedings of the IEEE International Radar Conference, Arlington, VA, USA, May 9 - 12, 2005.

*[10] R. Srinivasan and M. Rangaswamy, "Importance sampling for characterizing STAP detectors", IEEE Trans. AES, January 2007.

*[11] L. Anitori, R. Srinivasan and M. Rangaswamy, "Importance sampling for NMF class of STAP detectors", CIE Radar 2006, Oct. 16 - 19, Shanghai, China.

*[12] L. Anitori, R. Srinivasan and M. Rangaswamy, "Envelope-law and geometric-mean STAP detection" , *in preparation.*

[13] K. R. Gerlach, "New Results in Importance Sampling," *IEEE Trans. Aerospace and Electronic Systems*, vol. 35, No. 3, July 1999, pp 917-925.

[14] D. L. Stadelman, D. D. Weiner, and A. D. Keckler, "Efficient determination of thresholds via importance sampling for Monte Carlo evaluation of radar performance in non-Gaussian clutter" *Proceedings of the IEEE Radar Conference*, Long Beach, CA, April 2002, pp. 272-277.

[15] M. Rangaswamy, J. H. Michels, and B. Himed, "Performance of Parametric and Covariance Based STAP Tests in Compound-Gaussian Clutter", *Digital Signal Processing*, vol. 12, pp. 307-307, 2002.

[16] M. Rangaswamy, F. C. Lin, K. R. Gerlach, "Robust adaptive signal processing methods for heterogeneous radar clutter scenarios", *Signal Processing*, vol. 84, NO. 9, pp. 1653-1665, September 2004.

[17] M. Rangaswamy, F. C. Lin, "Normalized Adaptive Matched Filter-A Low Rank Approach", *IEEE Symposium on signal Processing and information technology*,pp. 182-185, December 2003.

[18] I. Kirsteins, D. Tufts, "Adaptive detection using a low rank approximation to a data matrix ", *IEEE Trans. Aerospace and Electronic Systems*, AES-30, NO. 1, pp. 55-67, January 1994.

[19] P. Chen, M. C. Wicks, R. S. Adve, "Development os a statistical procedure for detecting the number of signals in a radar measurement ", *IEE Proc. Radar, Sonar Navig.*, Vol. 148, NO.4, pp. 219-226, August 2001.

[20] M. S. Srivastava, C. G. Khatri, "An introduction to multivariate statistics", Elsevier North Holland, 1979.

* *These papers are outcomes of this research effort.*

Importance sampling for characterizing STAP detectors

Rajan Srinivasan and Muralidhar Rangaswamy

***Abstract*— This paper describes the development of adaptive importance sampling techniques for estimating false alarm probabilities of detectors that use space-time adaptive processing (STAP) algorithms. Fast simulation using importance sampling methods has been notably successful in the study of conventional constant false alarm rate (CFAR) radar detectors, and in several other applications. The principal objectives here are to examine the viability of using these methods for STAP detectors, develop them into powerful analysis and design algorithms and, in the long term, use them for synthesizing novel detection structures. The adaptive matched filter (AMF) detector has been analyzed successfully using fast simulation. Of two biasing methods considered, one is implemented and shown to yield good results. The important problem of detector threshold determination is also addressed, with matching outcome. As an illustration of the power of these methods, two variants of the square-law AMF detector that are thought to be robust under heterogeneous clutter conditions have also been successfully investigated. These are the envelope-law and geometric-mean STAP detectors. Their CFAR property is established and performance evaluated. It turns out the variants have detection performances better than those of the AMF detector for training data contaminated by interferers. In summary, the work reported here paves the way for development of advanced estimation techniques that can facilitate design of powerful and robust detection algorithms.**

Keywords STAP, AMF, Importance Sampling (IS), CFAR, GM-STAP, E-AMF, Homogeneous clutter, Complex Gaussian PDF.

I. INTRODUCTION

Estimation of false alarm probabilities of detection algorithms that employ space-time processing is examined here using forced Monte Carlo (MC) or importance sampling (IS) simulation. Space-time adaptive processing algorithms are of much importance for radar detection, [1], [2]. They are notoriously intensive from a computational point of view, with the more advanced (and robust) ones being also analytically difficult to quantify. Therefore, it is appropriate to attempt to develop fast simulation methods that could be used in their analysis and design.

The work of the first author was supported by the European Office of Aerospace Research and Development, London, UK, in collaboration with the Air Force Office of Scientific Research under Award No. FA8655-04-1-3025. Portions of this paper were presented at the IEEE International Radar Conference, Washington, DC, USA, May 2005.

R. Srinivasan is with the Telecommunication Engineering Group, University of Twente, PO 217, 7500 AE Enschede, The Netherlands (email: r.srinivasan@ewi.utwente.nl).

M. Rangaswamy is with the Air Force Research Laboratory Sensors Directorate, Hanscom AFB, MA, USA (email:Muralidhar.Rangaswamy@hanscom.af.mil).

In the following we use lessons learnt from developing IS techniques for characterizing conventional CFAR detectors and describe an experiment in applying them to STAP detection. The starting point of this effort is the celebrated AMF detector derived in [2] and which represents the array version of the workhorse cell averaging (CA) CFAR detector for conventional radar signal processing algorithms. The false alarm probability (FAP) performance of the AMF detector is known in integral form under certain conditions and can be numerically computed to any desired accuracy. Thus it forms a suitable basis for validating our simulation experiments. Two specific IS methods (described in the sequel) are presented and the better (and also easier) one is implemented. As a demonstration of the power of IS methods, the envelope-law and geometric mean detectors are presented and analyzed using fast simulation. These are variants of the AMF detector and their FAPs are not known in analytical form. Their CFAR property is established and FAP behaviour characterized using IS. A brief comparison of detection performances of all 3 detectors is made. In the long term, our aim is to develop and apply these fast simulation techniques to modern STAP detection algorithms (see [3] - [5], and references therein). An important goal in this context is to devise IS biasing methods that result in simulation times that grow slowly with decreasing FAPs. This remains an open problem.

As well known now, IS is the chief simulation methodology for rare-event estimation. It is an enduring method that has had success in several areas of science and engineering, [6]. Briefly, IS works by biasing original probability distributions in ways that accelerate the occurrences of rare events, conducting simulations with these new distributions, and then compensating the obtained results for the changes made. The principal consequence is that unbiased probability estimates with low variances are obtained quickly. The main task in IS therefore is determination of good simulation distributions for an application. Simulations performed using such distributions can provide enormous speed-ups and, if applied successfully, simulation lengths needed to estimate very low probabilities can become (only) weakly dependent on the actual probabilities. When these probabilities satisfy a large deviations principle, then several asymptotic results are available for devising simulation distributions, [6]. For most detection applications however, it appears that adaptive methods which attempt to minimize error variances might be better suited. There have been some recent attempts in the literature, for example [7], [8], to apply IS for FAP estimation of CFAR detectors with varying degrees of success. Our work uses results developed in [9] and [10], with an adaptive implementation of the so

called g-method. Recent theoretical and application work on this method can be found in [11] and [12]. The groundwork for the present results was laid down in [15] which, to our knowledge, is the first article on the topic of using IS for studying STAP detection algorithms.

During the conduct of simulations reported herein, some issues concerning the adaptive IS algorithms used arise, and these are discussed briefly. The positive outcome of the methods used is that excellent match with numerical results is obtained. The succeeding sections provide a brief statement of the detector tests, how IS biasing can be performed to hasten false alarm events, description of the g-method which is a conditional IS technique developed originally for studying sums of random variables [9], how *inverse IS* can be used to estimate (and choose) detector thresholds, a methodology for using the fast adaptive algorithms, simulation results, and a conclusion.

II. STAP DETECTOR VARIANTS

In a radar system consisting of a linear array of N_s antenna elements a burst of N_t pulses is transmitted, resulting in a return of $N_sN_t = N$ samples in each range gate. The N complex samples are arranged in an $N \times 1$ column vector and there are $L+1$ such returns. The range gate return to be tested for the presence of target is called the primary data and is denoted by $\mathbf{x}$ while the remaining vectors are known as the training (or secondary) data, assumed to be free of target signal, and denoted by $\mathbf{x}(l)$, $l = 1, \ldots, L$. A target return is modelled as consisting of an $N \times 1$ steering vector $\mathbf{s}$ with an unknown complex amplitude in addition to clutter, interference, and noise. The primary and secondary data vectors are assumed to be jointly independent and zero-mean, spherically symmetric complex Gaussian, sharing the $N \times N$ covariance matrix $\mathbf{R} = E\{\mathbf{X}\mathbf{X}^\dagger\}$, where the superscript $\dagger$ denotes complex conjugate transpose.

A. The AMF (square-law) STAP detector

Under these assumptions the AMF detection test, as obtained in [2], is given by

$$\frac{|\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}|^2}{\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{s}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{1}$$

where

$$\widehat{\mathbf{R}} \equiv \frac{1}{L}\sum_{l=1}^{L} \mathbf{x}(l)\mathbf{x}(l)^\dagger$$

is the estimated covariance matrix of $\mathbf{x}$ based on the secondary data (also referred to as sample matrix), and η is a threshold used to set the FAP at some desired level. This test has the CFAR property. The FAP α of the test is known to be given by

$$\alpha = \frac{L!}{(L-N+1)!(N-2)!}\int_0^1 \frac{x^{L-N+1}(1-x)^{N-2}}{(1+\eta\, x/L)^{L-N+1}}\, dx \tag{2}$$

which can be used to numerically determine the threshold setting for a desired FAP. As shown in [2], the test in (1) can be rewritten as

$$|\,\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \eta\, \mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{s} = \frac{\eta}{L}\sum_{l=1}^{L} |\,\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}(l)|^2 \tag{3}$$

This is in the form of a vector (or, array) version of the usual CA-CFAR test. The LHS is a square law detector, being the output of a matched filter (matched to the direction $\mathbf{s}$ in which the array is steered) for incoherent detection using the so-called sample matrix inversion (SMI) beamformer weights $\widehat{\mathbf{R}}^{-1}\mathbf{s}$. The RHS represents a cell averaging term. Further details on these issues can be found in the references mentioned above.

B. The AMF (envelope-law) STAP detector

In contrast to square-law detectors, it is known to radar engineers that envelope detectors possess some robustness properties in terms of detection performance when the training data is contaminated by outliers or inhomogeneities. Accordingly, the envelope-law STAP version of the AMF (abbreviated hereafter as E-AMF) detector is proposed here and its detection properties evaluated. It takes the form

$$|\,\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}| \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\eta_E}{L}\sum_{l=1}^{L} |\,\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}(l)| \tag{4}$$

where η_E denotes the threshold multiplier. An analytical expression or approximation for the FAP of this detector is not known.

C. The geometric-mean STAP detector

Also proposed here is the geometric-mean STAP version of the AMF detector (abbreviated as GM-STAP). Conventional CFAR detectors that calculate the geometric means of the range cells in the CFAR window (usually referred to as log-t detectors) are known to have robustness properties. The STAP variant[1] takes the form

$$|\,\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}| \underset{H_0}{\overset{H_1}{\gtrless}} \eta_G \left(\prod_{l=1}^{L} |\,\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}(l)|\right)^{1/L} \tag{5}$$

where η_G denotes the threshold multiplier. Note that the GM-based square- and envelope-law versions are identical (except for a trivial squaring of threshold multiplier). For the sake of completeness, the value of the multiplier as the number of training vectors $L \to \infty$ is calculated. In such a case, $\widehat{\mathbf{R}}^{-1} \xrightarrow{p} \mathbf{R}^{-1}$, $\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{X} \xrightarrow{D} \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{X}$, and $\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{X}(l) \xrightarrow{D} \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{X}(l)$ in the absence of target. As $\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{X}$ and $\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{X}(l)$ are i.i.d. and distributed as $\mathcal{CN}_1(0, \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s})$, it follows that the squared envelopes $|\,\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{X}|^2$ and $|\,\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{X}(l)|^2$ are exponentially distributed,

[1] This test was actually suggested in [16] but its performance was not evaluated.

each with mean $\mathbf{s}^\dagger\mathbf{R}^{-1}\mathbf{s}$. Using convergence arguments based on continuity, [14], it is straightforward to show that the asymptotic FAP of the GM detector is $\exp(-\eta_G^2\exp(-\gamma))$ where γ is the Euler constant. Hence the (asymptotic) threshold η_G required to provide a FAP of 10^{-6} is 4.9605. A similar calculation can be easily made for the E-AMF detector.

Both the E-AMF and GM-STAP detectors have the CFAR property under the assumption of homogeneous Gaussian characterization as stated in the beginning of this section; their FAPs are independent of the true target-free covariance $\mathbf{R}$. This is shown in Appendix I.

III. FAP ESTIMATION USING IS

This section describes procedures to quickly estimate the FAPs and threshold multipliers of STAP detectors using IS. The standard (square-law) AMF detector is used for exposition. The detector variants are handled using parallel arguments.

The actual form of biasing considered here is simple scaling. While other techniques (superior in terms of simulation gains) can be devised, scaling is easy to implement and yields conservative but reasonable results.

A. Square-law AMF detection

Two strategies to estimate FAPs using IS are two-dimensional (2-d) biasing and the conditional g-method procedure, described in this section. The 2-d biasing scheme parallels the approach used in previous works on conventional CFAR algorithms, as cited in the introduction. It can be useful in situations wherein the more powerful g-method might be difficult to apply.

1) 2-d biasing: To estimate FAP using IS, we make the following observations. Suppose each complex sample of a secondary vector is scaled by a real (and positive) number $\theta^{1/2}$. This has the effect of scaling the covariance matrix estimate $\widehat{\mathbf{R}}$ by θ. Therefore, as far as the covariance estimate is concerned, both sides of the test in (3) remain unaffected by the scaling. However, each secondary vector being scaled by $\theta^{1/2}$ results in an additional scaling of the RHS by θ. Hence choosing θ less than unity will have the effect of compressing the density function of the random threshold of the test. Further, a scaling of each complex component of the primary vector by a real positive $a^{1/2}$ will achieve a scaling of the LHS of the test by a. Thus, choosing a larger and θ smaller than unity will achieve an increase in the frequency of occurrence of false alarm events during simulation. The IS optimization problem will be a two-parameter one.

Denoting by $\mathcal{A}$ the false alarm event in (3), the unbiased IS estimator in an i.i.d. (independent and identically distributed) simulation can be expressed as

$$\widehat{\alpha} = \frac{1}{K}\sum_{i=1}^{K}[1(\mathcal{A})\,W(\mathbf{X},\mathbf{X}_L;a,\theta)]^{(i)}; \quad \mathbf{X},\mathbf{X}_L \sim f_\star \tag{6}$$

where K is the length (or number of trials) of the IS simulation, $1(\cdot)$ denotes the indicator, $\mathbf{X}_L \equiv (\mathbf{X}(1),\ldots,\mathbf{X}(L))^T$, and the notation $\sim f_\star$ means that all random variables are drawn from biased distributions. The weighting function W is described below. In setting up the joint densities of $\mathbf{X}$ and $\mathbf{X}_L$, we use the fact that the FAP of the AMF has the CFAR property and is independent of the true covariance matrix $\mathbf{R}$. This is so under the assumption of Gaussian distributions for the data. In such a case, the simulation of the AMF test can be carried out for data possessing a diagonal covariance matrix $\mathbf{I}$, denoting the $N\times N$ identity matrix. Therefore, primary and secondary data can be generated as complex vectors with independent components. The unbiased joint densities are

$$f(\mathbf{x}) = \frac{e^{-\mathbf{x}^\dagger\mathbf{x}}}{\pi^N} \text{ and } f(\mathbf{x}_L) = \frac{e^{-\sum_1^L \mathbf{x}(l)^\dagger\mathbf{x}(l)}}{\pi^{LN}}$$

so that

$$f(\mathbf{x},\mathbf{x}_L) = \frac{e^{-\mathbf{x}^\dagger\mathbf{x}-\sum_1^L \mathbf{x}(l)^\dagger\mathbf{x}(l)}}{\pi^{(L+1)N}}$$

With scaling performed as described above, the biased joint density takes the form

$$f_\star(\mathbf{x},\mathbf{x}_L) = \frac{e^{-\frac{1}{a}\mathbf{x}^\dagger\mathbf{x}-\frac{1}{\theta}\sum_1^L \mathbf{x}(l)^\dagger\mathbf{x}(l)}}{\pi^{(L+1)N}a^N\theta^{LN}}$$

resulting in the weighting function

$$\begin{aligned} W(\mathbf{X},\mathbf{X}_L;a,\theta) &\equiv \frac{f(\mathbf{x},\mathbf{x}_L)}{f_\star(\mathbf{x},\mathbf{x}_L)} \\ &= Ca^N\theta^{LN}e^{A/a}e^{B/\theta} \end{aligned} \tag{7}$$

where

$$A \equiv \mathbf{x}^\dagger\mathbf{x}, \quad B \equiv \sum_1^L \mathbf{x}(l)^\dagger\mathbf{x}(l), \text{ and } C \equiv e^{-(A+B)} \tag{8}$$

The variance of the IS estimator $\widehat{\alpha}$ can be expressed as

$$\operatorname{var}\widehat{\alpha} = \frac{1}{K}[I(\boldsymbol{\nu}) - \alpha^2] \tag{9}$$

with $\boldsymbol{\nu}$ denoting the vector biasing parameter $(a,\theta)^T \in [1,\infty)\times(0,1]$. The quantity I is given by

$$\begin{aligned} I(\boldsymbol{\nu}) &= E_\star\{1(\mathcal{A})W^2(\mathbf{X},\mathbf{X}_L;\boldsymbol{\nu})\} \\ &= E\{1(\mathcal{A})W(\mathbf{X},\mathbf{X}_L;\boldsymbol{\nu})\} \end{aligned} \tag{10}$$

where the expectation $E_\star$ proceeds over biased distributions. Minimization of $\operatorname{var}\widehat{\alpha}$ with respect to the biasing parameters is equivalent to minimization of I and is described in Appendix II.

2) The g-method estimator: This method exploits knowledge of underlying (input) distributions more effectively, yielding a more powerful estimator. Additional advantages are that only a scalar parameter optimization problem needs to be tackled and the *inverse IS* problem (for threshold optimization or selection) can be easily solved. The FAP can be written as

$$\begin{aligned} \alpha &= P(\mathcal{A}) \\ &= E\{P(|\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}|^2 > \eta\,\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s} \mid \mathbf{X}_L, H_0)\} \\ &\triangleq E\{g(\mathbf{X}_L)\} \end{aligned} \tag{11}$$

where $g(\mathbf{X}_L)$ denotes the conditional probability in the second step. The conditioning implies that the covariance matrix estimate $\widehat{\mathbf{R}}$ is given. We proceed to estimate α using the form in the third step above.

With the conditioning in mind it is easy to show, assuming that $\mathbf{X}$ is rotationally invariant and Gaussian, that the random variable $\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X} \triangleq \mathbf{w}^\dagger\mathbf{X}$ is distributed as $\mathcal{CN}_1(0, \mathbf{w}^\dagger\mathbf{R}\mathbf{w})$ with independent real and imaginary components, and the weight vector $\mathbf{w} = \widehat{\mathbf{R}}^{-1}\mathbf{s}$. Hence the random variable $Y \triangleq |\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{X}|^2$ is exponential and has the density function

$$f(y|\mathbf{X}_L, H_0) = \frac{e^{-y/\mathbf{w}^\dagger\mathbf{R}\mathbf{w}}}{\mathbf{w}^\dagger\mathbf{R}\mathbf{w}}, \quad y \geq 0$$

Therefore

$$\begin{aligned} g(\mathbf{X}_L) &= P(Y \geq \eta\, \mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s} \mid \mathbf{X}_L, H_0) \\ &= e^{-\eta\, D} \end{aligned}$$

where

$$D \equiv \frac{\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}}{\mathbf{w}^\dagger\mathbf{R}\mathbf{w}}$$

Note that if $\widehat{\mathbf{R}} = \mathbf{R}$, then $g(\mathbf{X}_L) = e^{-\eta}$ and this is the FAP of the AMF when the covariance matrix is known. As discussed before, we run simulations with homogeneous data possessing an identity covariance matrix, that is, with $\mathbf{R} = \mathbf{I}$. The g-method IS estimator then takes the form

$$\begin{aligned} \widehat{\alpha}_g &= \frac{1}{K}\sum_{i=1}^{K}[g(\mathbf{X}_L)W(\mathbf{X}_L;\theta)]^{(i)} \\ &= \frac{1}{K}\sum_{i=1}^{K}[e^{-\eta\, D}W(\mathbf{X}_L;\theta)]^{(i)}; \quad \sim f_\star \end{aligned} \tag{12}$$

with D given by

$$\begin{aligned} D &= \frac{\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}}{\|\mathbf{w}\|^2} \\ &= \frac{\mathbf{s}^\dagger\widehat{\mathbf{R}}^{-1}\mathbf{s}}{\mathbf{s}^\dagger(\widehat{\mathbf{R}}^{-1})^2\mathbf{s}} \end{aligned} \tag{13}$$

Now, choosing the (single) biasing parameter $\theta < 1$ produces a decrease in D, thereby causing a higher frequency of occurrence of the false alarm event or, more appropriately in this case, a larger value of the g-function. Note that use of the g-method obviates the need to bias primary data vectors. Determination of a good value of θ proceeds as described in Appendix II. The weighting function is simply

$$W(\mathbf{x}_L;\theta) = \theta^{LN}e^{-(1-1/\theta)B} \tag{14}$$

which can be deduced from (7) by setting $a = 1$. The variance of this estimator is given by

$$\text{var}\,\widehat{\alpha}_g = \frac{1}{K}[I_g(\theta) - \alpha^2] \tag{15}$$

where

$$\begin{aligned} I_g(\theta) &\equiv E_\star\{g^2(\mathbf{X}_L)W^2(\mathbf{X}_L;\theta)\} \\ &= E\{g^2(\mathbf{X}_L)W(\mathbf{X}_L;\theta)\} \end{aligned} \tag{16}$$

The minimization of (an estimate of) I_g with respect to the biasing parameter θ is carried out using the recursion

$$\theta_{m+1} = \theta_m - \delta_\theta\frac{\widehat{I}'_g(\theta_m)}{\widehat{I}''_g(\theta_m)}, \quad m = 1, 2, \ldots \tag{17}$$

which is just a one-dimensional version of (35) in Appendix II. The reader is again referred to this appendix for definitions of quantities in the summands of the following estimators for I_g and its derivatives (with respect to θ) that are used in (17). These are respectively given by

$$\begin{aligned} \widehat{I}_g(\theta) &= \frac{1}{K}\sum_{i=1}^{K}[g^2(\mathbf{X}_L)W^2(\mathbf{X}_L;\theta)]^{(i)} \\ \widehat{I}'_g(\theta) &= \frac{1}{K}\sum_{i=1}^{K}[g^2(\mathbf{X}_L)W(\mathbf{X}_L;\theta)W_\theta(\mathbf{X}_L;\theta)]^{(i)} \\ \widehat{I}''_g(\theta) &= \frac{1}{K}\sum_{i=1}^{K}[g^2(\mathbf{X}_L)W(\mathbf{X}_L;\theta)W_{\theta\theta}(\mathbf{X}_L;\theta)]^{(i)} \end{aligned} \tag{18}$$

where $\mathbf{X}_L \sim f_\star$ in all 3 estimators above.

3) Threshold determination: The converse problem, namely that of finding by fast simulation the value of detector threshold η satisfying a prescribed FAP, is an important task that can be readily carried out using the estimator of (12). The genesis of the method lies in the so called *inverse IS* problem first formulated and solved in [9]. The solution is to minimize a "squared performance error" stochastic objective function

$$J(\eta) = [\widehat{\alpha}_g(\eta) - \alpha_o]^2$$

where α_o is a desired FAP. A typical example is shown in Figure 1 for an E-AMF detector. All detection algorithms that involve a threshold crossing will possess objective functions that have the general behaviour shown for $\alpha_o < 1$, assuming of course that the FAP estimate is (in a stochastic sense) monotone in the threshold η. Minimization of J with respect to η is carried out by the algorithm

$$\eta_{m+1} = \eta_m + \delta_\eta\frac{\alpha_o - \widehat{\alpha}_g(\eta_m)}{\widehat{\alpha'_g}(\eta_m)}, \quad m = 1, 2, \ldots \tag{19}$$

where δ_η is a step-size parameter and the derivative estimator in the denominator is given by

$$\widehat{\alpha'_g}(\eta_m) = -\frac{1}{K}\sum_{i=1}^{K}[D\, e^{-\eta D}\, W(\mathbf{X}_L;\theta)]^{(i)}; \quad \sim f_\star \tag{20}$$

with the prime indicating derivative with respect to η. Note that this derivative estimator actually estimates (negative of) the probability density function of the AMF statistic on the left hand side of (1) under the hypothesis that target is absent. The threshold-finding algorithm of (19) is a key component of the fast simulation methodology.

4) Simulation gain: A useful measure of the effectiveness of any IS procedure is the simulation gain Γ. It is the ratio of simulation lengths required by conventional MC and IS estimators to achieve the same error variance. The variance of an MC estimator ($W = 1$) is given by $(\alpha - \alpha^2)/K_{\text{MC}}$ for a simulation length K_{MC}. Hence, using (9), the IS gain for the estimator of (6) is

$$\Gamma = \frac{K_{\text{MC}}}{K} = \frac{\alpha - \alpha^2}{I(\boldsymbol{\nu}) - \alpha^2} \tag{21}$$

with $I(\boldsymbol{\nu})$ defined in (10). Similarly, using (15), the g-method estimator of (12) has gain Γ_g over MC estimation given by

$$\Gamma_g = \frac{\alpha - \alpha^2}{I_g(\theta) - \alpha^2} \tag{22}$$

where $I_g(\theta)$ is given by (16). Note that setting $\theta = 1$ in the above provides gain of the g-method estimator *without IS*. The estimator always has a smaller variance and consequently, $\Gamma_g > \Gamma$. Comparing (21) and (22), it is sufficient to show that $I_g < I$ to prove this. This is done in Appendix III.

As seen in the sequel, estimation of gain during simulation plays an important role in mechanizing an IS methodology that allows rare events to be studied quickly. In particular, Γ_g is estimated as

$$\widehat{\Gamma}_g = \frac{\widehat{\alpha}_g - \widehat{\alpha}_g^2}{\widehat{I}_g(\theta) - \widehat{\alpha}_g^2} \tag{23}$$

wherein the required estimates are given in (12) and (18). The denominator of the above equation is related to an estimate of the IS variance, denoted as $\widehat{\text{var}}\,\widehat{\alpha}_g$. Substituting (12) and (18) into (23) and performing a little algebra yields

$$\begin{aligned} \widehat{\text{var}}\,\widehat{\alpha}_g &\equiv \frac{\widehat{I}_g(\theta) - \widehat{\alpha}_g^2}{K} \\ &= \frac{1}{2K^3} \sum_{\substack{i=1 \\ i \neq j}}^{K} \sum_{j=1}^{K} [y_i - y_j]^2 \\ &> 0 \quad \text{(with probability 1)} \end{aligned} \tag{24}$$

where $y_i \equiv [g(\mathbf{X}_L) W(\mathbf{X}_L; \theta)]^{(i)}$. The variance estimate is asymptotically unbiased and it is easy to show, using (15) and (24), that $\widehat{\text{var}}\,\widehat{\alpha}_g \xrightarrow{p} \text{var}\,\widehat{\alpha}_g \to 0$ as $K \to \infty$, for *any* fixed θ.

If the IS estimator of (6) is used instead of (12), then $y_i \equiv [1(\mathcal{A})\, W(\mathbf{X}, \mathbf{X}_L; a, \theta)]^{(i)}$ and the inequality in (24) will not be strict, leading to a non-zero probability of instability in gain estimation. One way of trying to prevent this from happening is to overbias the simulation, but this can result in underestimation of the false alarm probability. Such instabilities do not occur with the g-method estimator and this is one of its implementation advantages.

B. E-AMF and GM-STAP detection

False alarm probability IS estimators for these two detector variants are set up in completely parallel fashion, starting with the expression in (12). The only difference is in the definition of the quantity D (given in (13) for the AMF detector), while the threshold multipliers get squared. This follows by observing from (3), (4), and (5) that all 3 detectors studied here differ, primarily, on the right hand sides of their respective tests. Consequently, denoting by D_E and D_G respectively the D values for the two variants, this leads to

$$D_E \equiv \frac{\left(\frac{1}{L} \sum_{l=1}^{L} |\, \mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{x}(l) | \right)^2}{\mathbf{s}^\dagger (\widehat{\mathbf{R}}^{-1})^2 \mathbf{s}}$$

and

$$D_G \equiv \frac{\left(\prod_{l=1}^{L} |\, \mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{x}(l) |^2 \right)^{1/L}}{\mathbf{s}^\dagger (\widehat{\mathbf{R}}^{-1})^2 \mathbf{s}}$$

The corresponding FAP estimators are obtained by using the above definitions in (12) with $g(\mathbf{X}_L)$ replaced by $\exp(-\eta_E^2 D_E)$ and $\exp(-\eta_G^2 D_G)$ respectively. The threshold finding algorithm in (19) is altered accordingly.

IV. Mechanizing the IS algorithm

A. A methodology

The various issues of the preceding sections are summarized in an IS simulation methodology that outlines the principal steps required for implementation of the adaptive algorithms. It is a cautious methodology which has been used in previous works on applications of IS.

Of interest is the extreme but realistic (and often encountered) situation where we have no knowledge *whatsoever* of the FAP α of a detection algorithm for a given value of threshold η. In such a case, referring to the objective function of Figure 1, the basic idea of fast adaptive simulation is to travel down the curve *from its left side*. An initial (α_o, η_o) pair is needed. It is easily obtained using conventional MC simulation for a high value of α_o, say 0.1 (and a correspondingly low value of η_o), using $K = 100/\alpha_o = 1000$ trials according to the well known MC thumb rule. This can be accomplished quickly with a few experiments. The (α_o, η_o) pair (whose accuracy need not be very high) provides the starting values for the IS procedure. It begins by forming the estimate $\widehat{I}_g(\theta)$ at the threshold η_o (using 1000 trials) and locating its minimum (possibly graphically) as a function of the biasing parameter θ. It is advisable to search for the optimum bias, denoted as θ_o, starting from $\theta = 1$ to avoid locating an overbiasing minimizer. This leads, via (23), to an estimate $\widehat{\Gamma}$ of the maximum IS gain available (with a possible correction for α) and results in the quadruplet $(\alpha_o, \eta_o, \theta_o, \widehat{\Gamma}(\theta_o))$. At this stage the IS adaptations are implemented. The threshold-finding algorithm of (19) is implemented with initial value η_o for a new prespecified α_o slightly smaller than the initial α_o and using a conservative number of $K = 100/(\alpha_o \widehat{\Gamma}(\theta_o))$ for the IS trials. It is conservative because we know that simulation gain increases with decreasing rare-event probability and hence this number is guaranteed to provide better than the rule-of-thumb accuracy at the new α_o. Note that most of the IS gain can be leveraged if the FAP decrements are kept small. The biasing algorithm of (17) is implemented simultaneously using θ_o as initial value. At the end of the adaptations the resulting gain is estimated for updating K and a new quadruplet is obtained. This is continued until we have a complete characterization of false-alarm behaviour of the detector down to the desired α_o. The procedure is summarized below.

1) Implementation: Define the set $\{\alpha_o^{(p)}\}_1^P$ such that $0.1 = \alpha_o^{(1)} > \alpha_o^{(2)} > \cdots > \alpha_o^{(P)} = 10^{-7}$, for example.

Pre-processing:

Step 1. $p = 1$. Use 1000 conventional MC trials to obtain $(\alpha_o^{(1)}, \eta_o^{(1)})$ pair.

Step 2. Find $\theta_o^{(p)} = \arg\min_\theta \widehat{I}_g(\theta)$ and calculate $\widehat{\Gamma}^{(p)}(\theta_o^{(p)})$ using (23).

Adaptive IS:

Step 3. $p = p + 1$

Step 4. Let $m = 1$. Set $\theta_m^{(p)} = \theta_o^{(p-1)}$, $\eta_m^{(p)} = \eta_o^{(p-1)}$ and $K = 100/(\alpha_o^{(p)}\widehat{\Gamma}^{(p-1)}(\theta_o^{(p-1)}))$.
Step 5. Generate K secondary vectors and compute B and D using (8) and (13).
Step 6. Implement (12), (18) using biased secondary vectors with parameter $\theta_m^{(p)}$.
Step 7. Compute $\theta_{m+1}^{(p)}$ and $\eta_{m+1}^{(p)}$ using (17) and (19); if both algorithms have converged, set $\theta_o^{(p)} = \theta_{m+1}^{(p)}$, $\eta_o^{(p)} = \eta_{m+1}^{(p)}$, $\widehat{\Gamma}^{(p)}(\theta_o^{(p)}) = \widehat{\Gamma}^{(p)}(\theta_{m+1}^{(p)})$ using (23), and go to step 3 or stop if $p = P$; otherwise go to step 5 with $m = m + 1$.

At the end of simulations we have the P quadruplets $\{\alpha_o^{(p)}, \eta_o^{(p)}, \theta_o^{(p)}, \widehat{\Gamma}^{(p)}(\theta_o^{(p)})\}_1^P$.

B. Complexity reduction for IS adaptations

The methodology is simple enough but certain observations can be made which lead to an artifice that further reduces computational effort to a considerable extent. In rare-event simulation of highly reliable systems that involve complex signal processing operations, two factors contribute to simulation time. The first is the rare event itself that is under study; this is handled by suitable IS biasing techniques. The second factor is the computational intensity of the signal processing of the system. In STAP detectors the chief processing burden is from inversion of large matrices. Several millions of MC trials with as many matrix inversions are needed to estimate low FAPs. Using an IS scheme the number of trials can be reduced appreciably, for a single rare-event probability estimation. However, there remains the important matter of executing the IS adaptations of (17) and (19), using this small number of IS trials for each iteration. These can be computationally burdensome if the recursions are to be carried out for several α_o values. This is where the artifice comes in. It turns out, in contrast to conventional MC simulation, that adaptations can be performed by *reusing* the unbiased random quantities generated for any one set of trials. Consequently it is not necessary to run truly randomized adaptive IS algorithms. The result will be a large savings in computational effort.

This idea is illustrated by rewriting the FAP estimator of (12) in the form

$$\widehat{\alpha}_g = \frac{1}{K}\sum_{i=1}^{K}\Big[\, e^{-\eta\,\theta\,D_{\text{U}}}\,\theta^{\,LN} e^{-(\theta-1)B_{\text{U}}}\Big]^{(i)}; \quad \sim f \qquad (25)$$

obtained by substituting (14) into (12), replacing D and B respectively by their *unbiased* versions denoted as D_{U} and B_{U}, and observing that whereas the estimator in (12) uses samples drawn from biased distributions $f_\star$, the mathematically equivalent estimator of (25) operates with only unbiased quantities, obtained via f. Biasing is handled by the (explicit) presence of θ. The number of trials remains unchanged, at K. The estimators required in (17) and (19) can be rewritten in a similar manner. It is a straightforward exercise and is omitted. Assuming then that the value of K is large enough for estimation of the target α_o (and this is guaranteed by the methodology described above), it follows that the set of K instances of D_{U} and B_{U} employed in (25) can be repeatedly used in the adaptations. With the complexity of frequent regeneration thus removed, these adaptations (such as minimization of $\widehat{I}_g(\theta)$ in (18)) can be extremely fast, often requiring only a few iterations. It is assumed that a sufficiently large pool of unbiased variables is pre-generated to accommodate K trials.

There are two related issues that are briefly discussed here in qualitative terms. In IS simulations carried out here (and elsewhere), it is tacitly assumed that the number of trials K is also large enough to accurately estimate $I_g(\theta)$. This quantity is admittedly not a rare-event probability and its estimator $\widehat{I}_g(\theta)$, given in (18), is clearly unbiased. Nevertheless, its accuracy should be studied in terms of its error variance, irrespective of whether we employ the above described reuse strategy or not. In this regard it must be pointed out that the biasing densities used for IS simulations are being chosen to minimize $I_g(\theta)$ (or its estimate $\widehat{I}_g(\theta)$), and not to estimate it accurately. Indeed, attempting to estimate $I_g(\theta)$ accurately would lead us to consider further biasing of the original biasing distribution. As probably known to IS researchers, the more powerful a biasing distribution is (in the sense of being "close" to the unrealizable optimal biasing density that estimates the rare-event probability exactly), the more "constant" will be the i.i.d. random variables y_i defined after (24). Interestingly this implies that even for small K, the estimated variance $\widehat{\text{var}}\,\widehat{\alpha}_g$ can become small for such biasing distributions. The latter observations accentuate the need to search for good simulation distributions. Unfortunately, it is beyond the remit of this paper to investigate the variance and rate of convergence of $\widehat{I}_g(\theta)$ and the dependencies of these on biasing distributions and number of trials K. Some remarks on this issue can be found in [16]. We have confirmed the legitimacy of the reuse strategy in several cases in two ways. Using an overkill, that is, employing a large number of IS trials for varying θ, it has been ascertained that the correct value of the optimum bias does not differ noticeably from the θ_o determined as above. Secondly, truly randomized optimization algorithms have been used, with the same conclusion.

The second issue concerns the ease with which the form of the estimator in (25) is obtained. The reason lies in the biasing technique used. Biasing in IS can be classed, for the purpose of this discussion, into two types. One kind derives biased quantities by direct parameterized transformations of unbiased variables, such as the scaling employed here. The second is just the complementary class, where such transformations do not exist. Examples of the latter can include new distributions for conducting simulations and some instances of distributions based on large deviations theory arguments. Within the first class, only the subset of transformations that are *many-to-one* will allow placing the IS estimator in a form such as that of (25).

V. Simulation results

Implementation results of the fast simulation procedures using the g-method are described here. A typical example of evolution of the threshold-finding algorithm is shown in Fig. 2 for the square-law AMF detector with $L = 704$ training

vectors and space-time product $N = 352$. Clearly, about 20 recursions of (19) appear sufficient for convergence. Note also from this figure that the "converged" value of the threshold for an α_o acts as the initial value for the next lower α_o, as described in Section IV-A. Although not shown here, similar estimation results have been obtained for the E-AMF and GM-STAP detectors. The results for the AMF detector almost "coincide" with numerical computations of (2). As analytical expressions or approximations for the FAPs are not available, such a comparison is not possible for the variants. Threshold multiplier values for the 3 detectors are summarized in Fig. 3 for $L = 128$ and $N = 64$. For the E-AMF and GM-STAP detectors the interpolations

$$\eta_E = 0.0095\,x^3 - 0.1713\,x^2 + 1.8869\,x + 1.6873$$
$$\eta_G = 0.0111\,x^3 - 0.1985\,x^2 + 2.2351\,x + 1.9899$$

for $L = 128$ and $N = 64$, and

$$\eta_E = 0.01\,x^3 - 0.1865\,x^2 + 1.8822\,x + 1.7255$$
$$\eta_G = 0.0118\,x^3 - 0.2198\,x^2 + 2.2292\,x + 2.0408$$

for $L = 704$ and $N = 352$, obtained from inverse IS results where $x = -\log_{10}\alpha$, can be used for estimating multipliers for FAPs in the range 10^{-1} to 10^{-7}.

Performance graphs of the IS algorithms are in Figures 4 and 5, which show the estimated IS gain Γ obtained over MC simulation and the number of trials required respectively. The number of IS trials, although appreciably reduced in comparison with MC, increases rapidly with decrease in false alarm probability. From these two figures it can be observed that, for a FAP of 10^{-6}, the square-law AMF with $L = 128$ and $N = 64$ requires 1500 IS trials whereas only 100 trials are needed for $L = 704$ and $N = 352$. This can be attributed to the fact that D and B, the only random quantities that appear in (12), have more concentrated density functions in the latter case, thus causing the IS biasing to be more effective during simulation.

Detection probabilities for all three detectors have been estimated and compared at a FAP of 10^{-6}. For homogeneous clutter backgrounds and non-fluctuating (Swerling 0) and fluctuating (Swerling 1) targets, the results are in Fig. 6. The SNR loss of the two variants compared to the AMF detector is very slight, the maximum being 0.3 db for a Swerling 0 target. Performance degradations for training data contaminated by nonhomogeneities consisting of interfering targets that are assumed to have the same Doppler-angle properties and characteristics as the primary target are shown in Fig. 7 for two Swerling 0 and Swerling 1 targets. Each interferer is assumed to have the same steering vector and power as the primary target. Similar results (not shown here) have been obtained for different numbers of interferers. It is evident that the GM-STAP detector is most robust in the presence of interferers, enjoying (in some cases) several decibels of power advantage over the square-law AMF detector. Consequently, its FAP performance in these situations (though not evaluated here) should also be relatively robust.

A. Comments

Despite the accuracy of our results, biasing by scaling has not done as well for the variants as for the AMF detector. Furthermore, even for the latter, IS performance should be improved by a better choice of biasing scheme in order to achieve the goal of having the required number of trials to be constant, or at least increasing very slowly, with decrease of FAP. This is certainly a matter for further investigation that may pay dividends while examining other detector configurations. Scaling was used in this work in the interest of expediency. Some details regarding the behaviour of the scaling parameter θ are available in [15]. Briefly, it is close to unity and has a small spread over the range of FAPs considered. This is due to the shape of the density functions of B and D.

VI. Conclusion

A humble inroad into the use of adaptive IS algorithms to characterize a STAP detector has been made. The AMF detector was used for validation and results have been good. The chief reasons for this are that we were able to invoke the g-method and inverse IS, use a biasing strategy that could be easily optimized adaptively, and find a way around the difficult task of inverting large matrices several times during simulations. As a small demonstration of the potential of IS, two AMF detector variants that are known to be relatively obdurate to mathematical analysis have been suggested and characterized. Sufficient numerical results have been provided here in order to motivate interested researchers to confirm our findings and possibly carry out their own investigations into the subject. There are some technical issues regarding IS estimators that remain to be settled; however, the development of better biasing methods is an important matter. In any case, our hope is that application of these fast simulation techniques to more advanced STAP configurations will also meet with success. But this remains to be seen as we are certainly not in position to predict what subtleties (and difficulties) these other detection algorithms can throw up. It is clear that IS is still in its infancy, especially insofar as its use for characterizing modern detection algorithms is concerned.

Appendix I
CFAR property

An invariance property is established that can be used to show that certain STAP detection algorithms have FAPs that do not depend on the data covariance $\mathbf{R}$. The proposition given here follows quite simply from arguments contained in the exposition of the generalized likelihood ratio STAP detector test (Kelly's GLRT) given in [1]. They are outlined here for convenience of the reader[2]. Assume, as before, that the primary and training data vectors have the same covariance. Consider the variables

$$G \equiv \mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x} \quad \text{and} \quad G(l) \equiv \mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1}\mathbf{x}(l) \tag{26}$$

[2]It would be helpful for the reader to refer to [1] (see also [2]). We have used several results from this now classic paper and have attempted to maintain the same notation.

for $l = 1, \ldots, L$. Using the transformations $\mathbf{u} = \mathbf{R}^{-1/2}\mathbf{s}$, $\mathbf{y} = \mathbf{R}^{-1/2}\mathbf{x}$, and $\mathbf{y}(l) = \mathbf{R}^{-1/2}\mathbf{x}(l)$, leads to

$$G = \mathbf{u}^\dagger \widetilde{\mathbf{R}}^{-1}\mathbf{y} \quad \text{and} \quad G(l) = \mathbf{u}^\dagger \widetilde{\mathbf{R}}^{-1}\mathbf{y}(l) \tag{27}$$

where $\widetilde{\mathbf{R}} \equiv \mathbf{R}^{-1/2}\widehat{\mathbf{R}}\mathbf{R}^{-1/2}$. The whitened vectors $\mathbf{Y}$ and $\mathbf{Y}(l)$ are both distributed $\mathcal{CN}_N(0, \mathbf{I})$. It turns out that $\widetilde{\mathbf{R}}$ has the complex Wishart distribution $\mathcal{CW}(L, N; \frac{1}{L}\mathbf{I})$, [13]. Further, a unitary transformation $\mathbf{U}$ can be found which rotates the new signal vector $\mathbf{u}$ into an elementary vector $\mathbf{e}$ as $d\mathbf{e} = \mathbf{U}^\dagger\mathbf{u}$, such that $\mathbf{e} = [1, 0, \ldots, 0]^\dagger$ and where $d^2 = \mathbf{u}^\dagger\mathbf{u} = \mathbf{s}^\dagger\mathbf{R}^{-1}\mathbf{s}$. The first column of $\mathbf{U}$ is the new signal vector $\mathbf{u}/d$. The remaining columns comprise an orthonormal basis determined, for example, by a Gram-Schmidt procedure. Let $\mathbf{z} = \mathbf{U}^\dagger\mathbf{y}$ and $\mathbf{z}(l) = \mathbf{U}^\dagger\mathbf{y}(l)$. Applying these to (27) yields the variables

$$G = \frac{d}{L}\,\mathbf{e}^\dagger\mathcal{S}^{-1}\mathbf{z} \quad \text{and} \quad G(l) = \frac{d}{L}\,\mathbf{e}^\dagger\mathcal{S}^{-1}\mathbf{z}(l) \tag{28}$$

where $\mathcal{S} \equiv L\,\mathbf{U}^\dagger\widetilde{\mathbf{R}}\mathbf{U}$. While $\mathbf{Z}$ and $\mathbf{Z}(l)$ are distributed as $\mathcal{CN}_N(0, \mathbf{I})$ and are independent, $\mathcal{S}$ has the distribution $\mathcal{CW}(L, N; \mathbf{I})$. The vectors $\mathbf{z}$ and $\mathbf{z}(l)$ are decomposed as $\mathbf{z} = [\mathbf{z}_{A} \;\; \mathbf{z}_{B}]^T$ and $\mathbf{z}(l) = [\mathbf{z}_{A}(l) \;\; \mathbf{z}_{B}(l)]^T$ where the A components are scalar and B components $(N-1)$-vector. Correspondingly, $\mathcal{S}$ is decomposed as

$$\mathcal{S} = \sum_{l=1}^{L}\mathbf{z}(l)\mathbf{z}(l)^\dagger = \begin{bmatrix}\mathcal{S}_{AA} & \mathcal{S}_{AB}\\ \mathcal{S}_{BA} & \mathcal{S}_{BB}\end{bmatrix} \tag{29}$$

so that $\mathcal{S}_{AB} = \sum_{l=1}^{L}\mathbf{z}_{A}(l)\mathbf{z}_{B}(l)^\dagger$, $\mathcal{S}_{BB} = \sum_{l=1}^{L}\mathbf{z}_{B}(l)\mathbf{z}_{B}(l)^\dagger$ and so on. Also

$$\mathbf{e}^\dagger\mathcal{S}^{-1} = \begin{bmatrix}\mathcal{P}_{AA} & \mathcal{P}_{AB}\end{bmatrix} \tag{30}$$

where

$$\begin{aligned}\mathcal{P}_{AA} &\equiv \left(\mathcal{S}_{AA} - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\mathcal{S}_{BA}\right)^{-1}\\ \mathcal{P}_{AB} &\equiv -\mathcal{P}_{AA}\mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\end{aligned} \tag{31}$$

which follow from the Frobenius relations for inversion of block matrices, [1]. Substituting (30) and (31) in (28) gives

$$G = \frac{d}{L}\,\mathcal{P}_{AA}\,y \quad \text{and} \quad G(l) = \frac{d}{L}\,\mathcal{P}_{AA}\,y(l) \tag{32}$$

where

$$\begin{aligned}y &\equiv \mathbf{z}_{A} - \sum_{l=1}^{L}\mathbf{z}_{A}(l)\mathbf{z}_{B}(l)^\dagger\,\mathcal{S}_{BB}^{-1}\mathbf{z}_{B}\\ y(l) &\equiv \mathbf{z}_{A}(l) - \sum_{i=1}^{L}\mathbf{z}_{A}(i)\mathbf{z}_{B}(i)^\dagger\,\mathcal{S}_{BB}^{-1}\mathbf{z}_{B}(l)\end{aligned} \tag{33}$$

Conditioned on the vectors $\mathbf{z}_{B}$ and $\mathbf{z}_{B}(l)$, it follows that the random variables Y and $Y(l)$ in (33) are (in the absence of target) zero mean Gaussian. With a little more algebra it can be shown, as in [1], that they are uncorrelated with variances

$$E_{B}\{|Y|^2\} \;=\; 1 + \mathbf{z}_{B}^\dagger\left(\sum_{l=1}^{L}\mathbf{z}_{B}(l)\mathbf{z}_{B}(l)^\dagger\right)^{-1}\mathbf{z}_{B}$$

and

$$E_{B}\{|Y(l)|^2\} = 1 - \mathbf{z}_{B}(l)^\dagger\left(\sum_{i=1}^{L}\mathbf{z}_{B}(i)\mathbf{z}_{B}(i)^\dagger\right)^{-1}\mathbf{z}_{B}(l)$$

for $l = 1, \ldots, L$, with E_{B} denoting conditional expectation. Further, the conditional covariance of the variables $Y(l)$ is given by

$$E_{B}\{Y(k)Y(n)^\star\} = -\,\mathbf{z}_{B}(n)^\dagger\left(\sum_{l=1}^{L}\mathbf{z}_{B}(l)\mathbf{z}_{B}(l)^\dagger\right)^{-1}\mathbf{z}_{B}(k)$$

for $k \neq n$.

Hence the set of conditionally jointly Gaussian zero mean random variables Y and $\{Y(l)\}_1^L$ have individual variances and covariances that are functions of the random vectors $\mathbf{z}_{B}$ and $\{\mathbf{z}_{B}(l)\}_1^L$. The latter are all jointly independent, each being distributed as $\mathcal{CN}_{N-1}(0, \mathbf{I})$. The probability of any event defined on the random variables Y and $\{Y(l)\}_1^L$ in (33) can thus be determined by performing an averaging operation over the distributions of $\mathbf{z}_{B}$ and $\{\mathbf{z}_{B}(l)\}_1^L$ and this probability will be independent of the data covariance $\mathbf{R}$. This statement is also true for the random variables G and $\{G(l)\}_1^L$ in (32) with the caveat that any constant scaling of these variables should leave the event unchanged. The preceding arguments therefore constitute proof of the following

Proposition: Any STAP detection algorithm that uses only the random variables G and $\{G(l)\}_1^L$ defined in (26) for its description such that the algorithm itself is unchanged by arbitrary but equal scaling of all these variables, has a FAP which is independent of the target-free data covariance $\mathbf{R}$.

It follows immediately from this that both the E-AMF and GM-STAP detectors have FAPs that are independent of the covariance matrix $\mathbf{R}$.

APPENDIX II
ADAPTIVE ALGORITHMS FOR 2-D BIASING

From the first line of (10), the I-function is estimated as

$$\widehat{I}(\boldsymbol{\nu}) = \frac{1}{K}\sum_{i=1}^{K}[1(\mathcal{A})\,W^2(\mathbf{X}, \mathbf{X}_L; \boldsymbol{\nu})]^{(i)}; \quad \sim f_\star \tag{34}$$

and its minimization can be carried out using the 2-dimensional adaptive algorithm

$$\boldsymbol{\nu}_{m+1} = \boldsymbol{\nu}_m - \delta\widehat{\mathbf{J}}_m^{-1}\widehat{\nabla}I(\boldsymbol{\nu}_m), \quad m = 1, 2, \ldots \tag{35}$$

Here, δ is a step-size parameter used to control convergence, and m is the index of iteration. This is a stochastic Newton recursion. It achieves minimization of $\widehat{I}$ by estimating a solution of

$$\widehat{\nabla}I(\boldsymbol{\nu}) \equiv (\widehat{I}_a \quad \widehat{I}_\theta)^T = \mathbf{0}$$

where $I_a \triangleq \partial I(\boldsymbol{\nu})/\partial a$ and $I_\theta \triangleq \partial I(\boldsymbol{\nu})/\partial\theta$. The estimate of the Jacobian $\mathbf{J}$ (which is the Hessian matrix of I) is given by

$$\widehat{\mathbf{J}} = \begin{pmatrix}\widehat{I}_{aa} & \widehat{I}_{a\theta}\\ \widehat{I}_{a\theta} & \widehat{I}_{\theta\theta}\end{pmatrix}$$

where $I_{xy} \equiv \partial I_x/\partial y$. Starting with the second line of (10), the various I-functions defined above can be obtained by using the notational equations

$$\begin{aligned} I_x &= E\{1(\mathcal{A})\,\partial W/\partial x\} \\ &= E_\star\{1(\mathcal{A}) W W_x\} \end{aligned}$$

and similarly

$$\begin{aligned} I_{xx} &= E_\star\{1(\mathcal{A}) W W_{xx}\} \\ I_{xy} &= E_\star\{1(\mathcal{A}) W W_{xy}\} \end{aligned}$$

with the corresponding weighting function derivatives easily calculated, using (7), as $W_x = \partial W/\partial x$, $W_{xy} = \partial^2 W/\partial x \partial y$, and so on. These I-functions can be estimated as in (34).

APPENDIX III
GAIN OF THE g-METHOD ESTIMATOR

The proof given here is in the context of STAP detectors and is a generalization of the one in [10] for conventional CFAR detection algorithms. A more complete generalization for estimators in an arbitrary rare-event setting is available in [11].

When no IS is used ($W = 1$), then $I_g = E\{g^2(\mathbf{X}_L)\} < E\{g(\mathbf{X}_L)\} = \alpha$. To show that $I_g < I \equiv I(\boldsymbol{\nu})$ (the biasing vector is omitted for convenience) with IS, some additional notation is useful. Let $U \equiv |\mathbf{s}^\dagger \widehat{\mathbf{R}}^{-1} \mathbf{X}|^2$ and let V denote the right hand side of any of the three STAP detectors, excluding the multiplier. We bear in mind that U is a function of $\mathbf{X}$ and $\mathbf{X}_L$, and V a function of $\mathbf{X}_L$. Then, in the absence of target, $\alpha = P(U \geq \eta\, V)$ and (10) can be written as

$$I = E\{1(U \geq \eta\, V) W(\mathbf{X}, \mathbf{X}_L)\}$$

with W defined in the first line of (7). For the g-method estimator, we can write

$$\begin{aligned} g(\mathbf{x}_L) &= P(U \geq \eta\, V \mid \mathbf{X}_L = \mathbf{x}_L) \\ &= E\{1(U \geq \eta\, V) \mid \mathbf{X}_L = \mathbf{x}_L\} \\ &= \int 1(u \geq \eta\, v) f(\mathbf{x}|\mathbf{x}_L)\, d\mathbf{x} \\ &= \int 1(u \geq \eta\, v)\, W_c(\mathbf{x}, \mathbf{x}_L)\, f_\star(\mathbf{x}|\mathbf{x}_L)\, d\mathbf{x} \end{aligned}$$

where $W_c(\mathbf{x}, \mathbf{x}_L) \equiv f(\mathbf{x}|\mathbf{x}_L)/f_\star(\mathbf{x}|\mathbf{x}_L)$. Then

$$\begin{aligned} g^2(\mathbf{x}_L) &\leq \int 1(u \geq \eta\, v)\, W_c^2(\mathbf{x}, \mathbf{x}_L)\, f_\star(\mathbf{x}|\mathbf{x}_L)\, d\mathbf{x} \\ &= \int 1(u \geq \eta\, v)\, W_c(\mathbf{x}, \mathbf{x}_L)\, f(\mathbf{x}|\mathbf{x}_L)\, d\mathbf{x} \end{aligned}$$

by Jensen's inequality. Therefore I_g, defined in (16), is given by

$$\begin{aligned} I_g &= \int g^2(\mathbf{x}_L)\, W(\mathbf{x}_L)\, f(\mathbf{x}_L)\, d\mathbf{x}_L \\ &\leq \int\int 1(u \geq \eta\, v)\, W_c(\mathbf{x}, \mathbf{x}_L)\, W(\mathbf{x}_L)\, f(\mathbf{x}|\mathbf{x}_L) \\ &\qquad \cdot f(\mathbf{x}_L)\, d\mathbf{x}\, d\mathbf{x}_L \\ &= \int\int 1(u \geq \eta\, v)\, W(\mathbf{x}, \mathbf{x}_L)\, f(\mathbf{x}, \mathbf{x}_L)\, d\mathbf{x}\, d\mathbf{x}_L \\ &= I. \end{aligned}$$

ACKNOWLEDGMENT

The authors are grateful to Dr. Michael Milligan of EOARD and Dr. Arje Nachman of AFOSR for their support and encouragement during the course of this work. Thanks are also due to Rusdha Muharar and Laura Anitori for software support and discussions.

REFERENCES

[1] E.J. Kelly, "An adaptive detection algorithm", *IEEE Trans. Aerospace and Electronic Systems*, AES-22, NO. 1, pp. 115-, March 1986.
[2] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerospace and Electronic Systems*, AES-28, NO. 1, pp. 208-216, January 1992.
[3] K. R. Gerlach, "Outlier resistant adaptive matched filtering," *IEEE Trans. Aerospace and Electronic Systems*, AES-38, NO. 3, pp. 885- 901, July 2002.
[4] J. H. Michels, M. Rangaswamy, and B. Himed,"Peformance of parametric and covariance based STAP tests in compound-Gaussian clutter," *Digital Signal Processing*, vol. 12, NOs. 2, 3, pp. 307-328, April-July 2002.
[5] M. Rangaswamy, "Statistical analysis of the non-homogeneity detector for non-Gaussian interference backgrounds," *IEEE Trans. Signal Processing*, vol. 53, NO. 6, pp. 2101-2111, June 2005.
[6] J. A. Bucklew, *Introduction to rare event simulation.* New York: Springer, 2004.
[7] K. R. Gerlach, "New Results in Importance Sampling", *IEEE Trans. Aerospace and Electronic Systems*, vol. 35, No. 3, July 1999, pp 917-925.
[8] D. L. Stadelman, D. D. Weiner, and A. D. Keckler, "Efficient determination of thresholds via importance sampling for Monte Carlo evaluation of radar performance in non-Gaussian clutter", *Proceedings of the IEEE Radar Conference*, Long Beach, CA, April 2002, pp. 272-277.
[9] R. Srinivasan, "Some results in importance sampling and an application to detection", *Signal Processing*, vol 65, Issue 1, pp 73-88, February 1998.
[10] R. Srinivasan, "Simulation of CFAR detection algorithms for arbitrary clutter distributions," *Proc. IEE, Radar, Sonar and Navigation*, vol. 147, February 2000, pp. 31-40.
[11] J. A. Bucklew, "Conditional importance samplng estimators", *IEEE Trans. Inform. Theory*, vol. 51, NO. 1, January 2005.
[12] G. V. Weinberg, "Estimation of false alarm probabilities in cell averaging constant false alarm rate detectors via Monte Carlo methods", *DSTO Systems Science Laboratory*, Report DSTO-TR-1624, Australia, November 2004.
[13] D. Maiwald and D. Kraus, "Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices", *IEE Proc. Radar, Sonar Navig.*, Vol. 147, NO.4, August 2000.
[14] P. Billingsley, *Convergence of Probability Measures*, John Wiley, 1999.
[15] R. Srinivasan and M. Rangaswamy, "Fast estimation of false alarm probabilities of STAP detectors - the AMF", Proceedings of the IEEE International Radar Conference, Arlington, VA, USA, May 9 - 12, 2005.
[16] R. Srinivasan, *Importance Sampling - Applications in Communications and Detection.* Berlin: Springer-Verlag, 2002.

LIST OF FIGURES

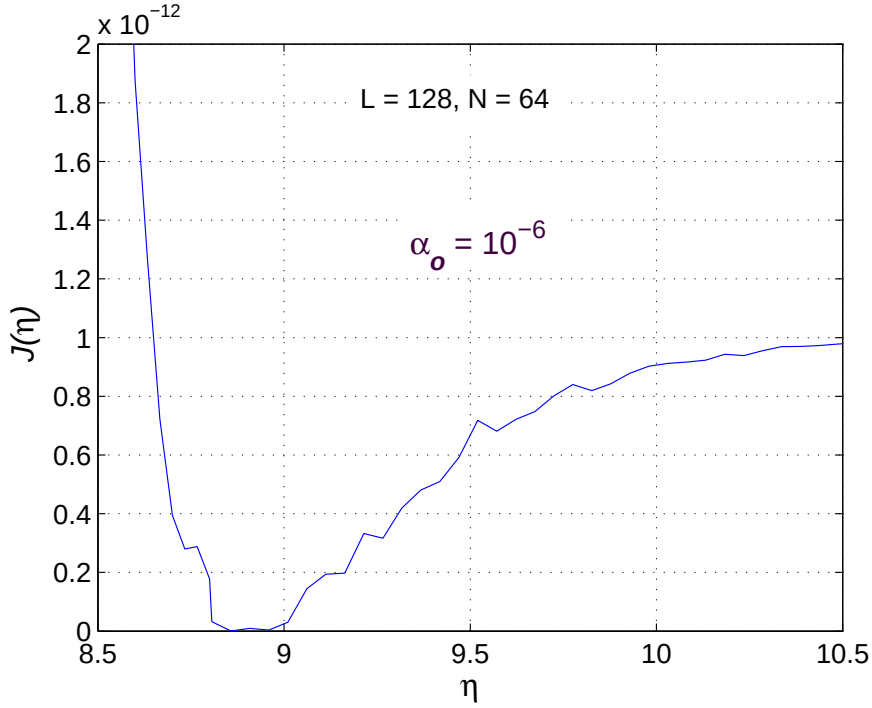

Fig. 1. Example objective function (for the E-AMF detector).

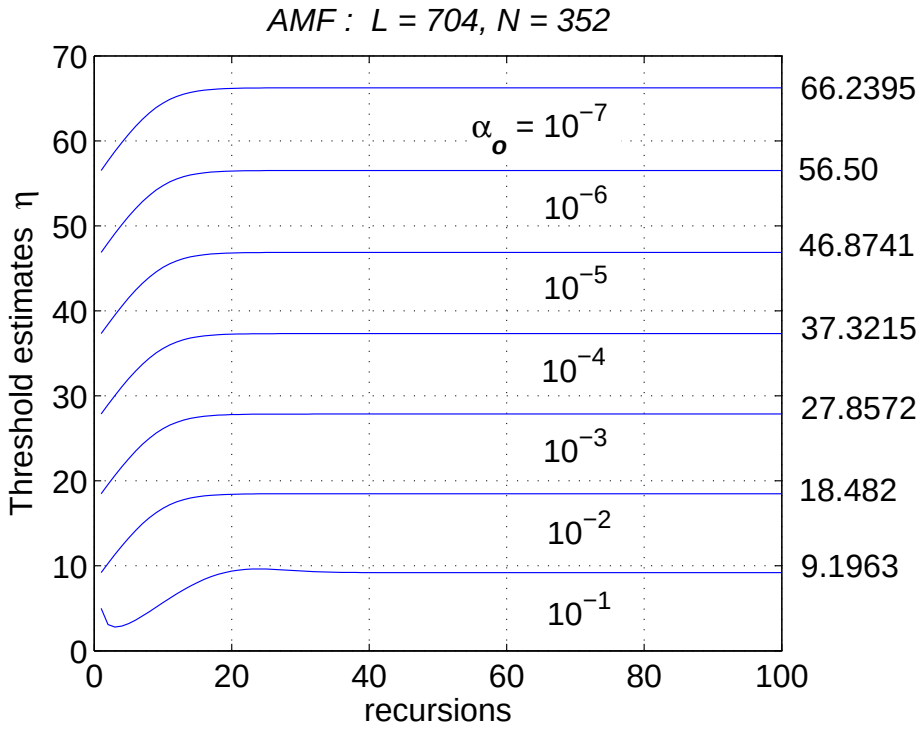

Fig. 2. Threshold optimization for AMF detector ($L = 704, N = 352$) using inverse IS.

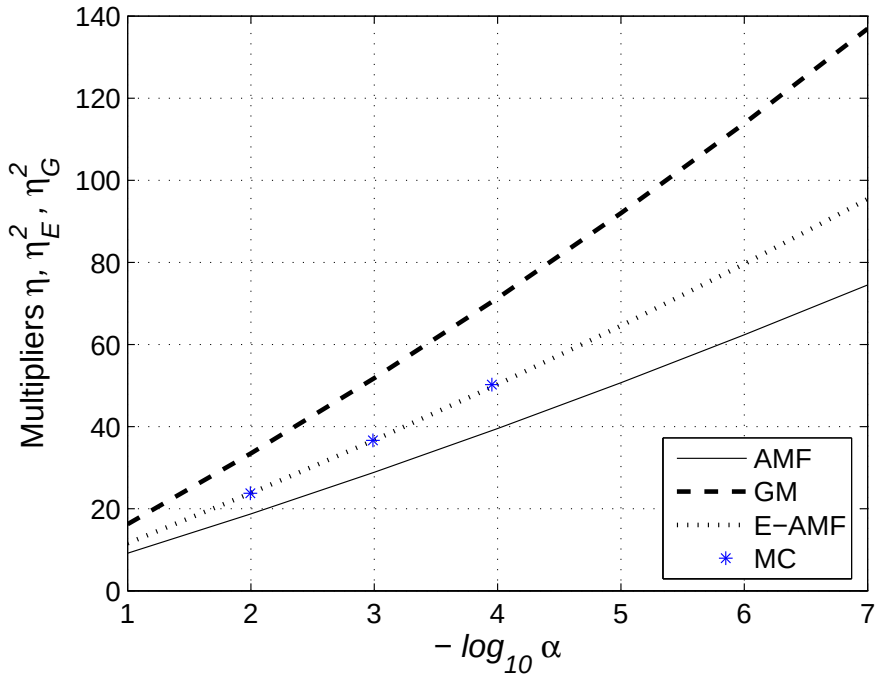

Fig. 3. Multipliers for $L = 128$ and $N = 64$. The starred dots indicate validations for the E-AMF detector using conventional MC simulations.

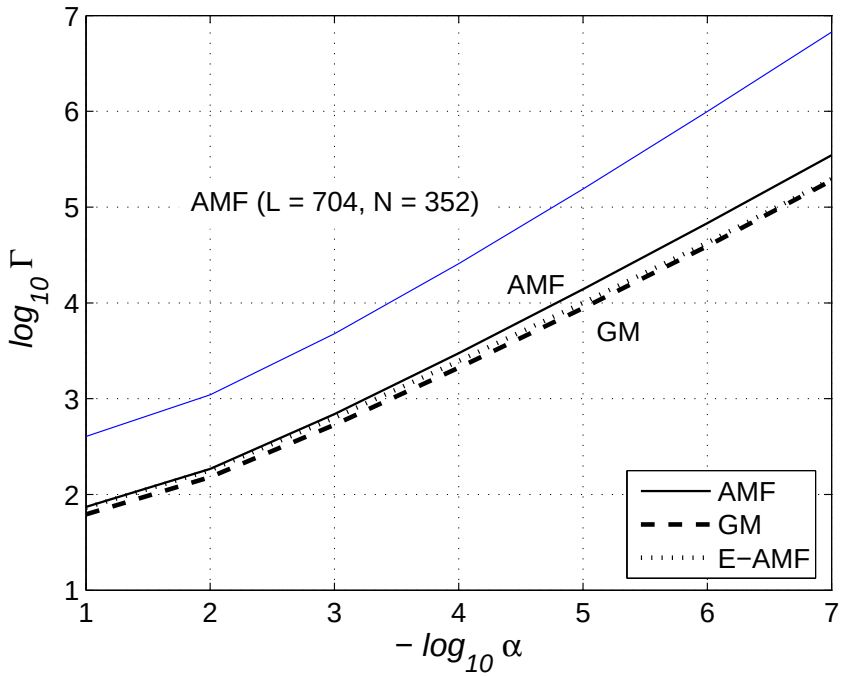

Fig. 4. IS gains. The set of 3 lower graphs correspond to $L = 128$ and $N = 64$.

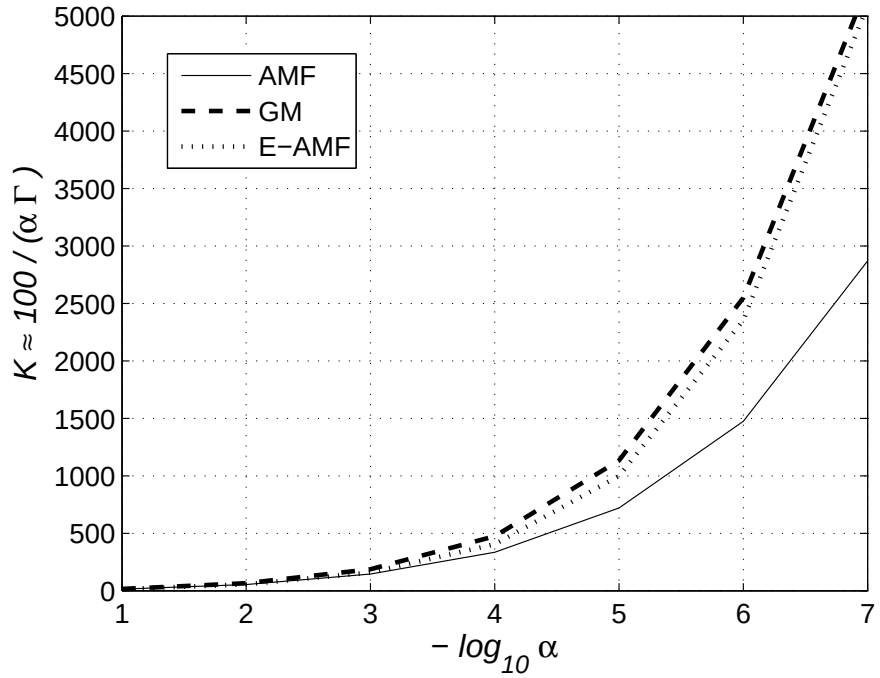

Fig. 5. Simulation lengths for $L = 128$ and $N = 64$.

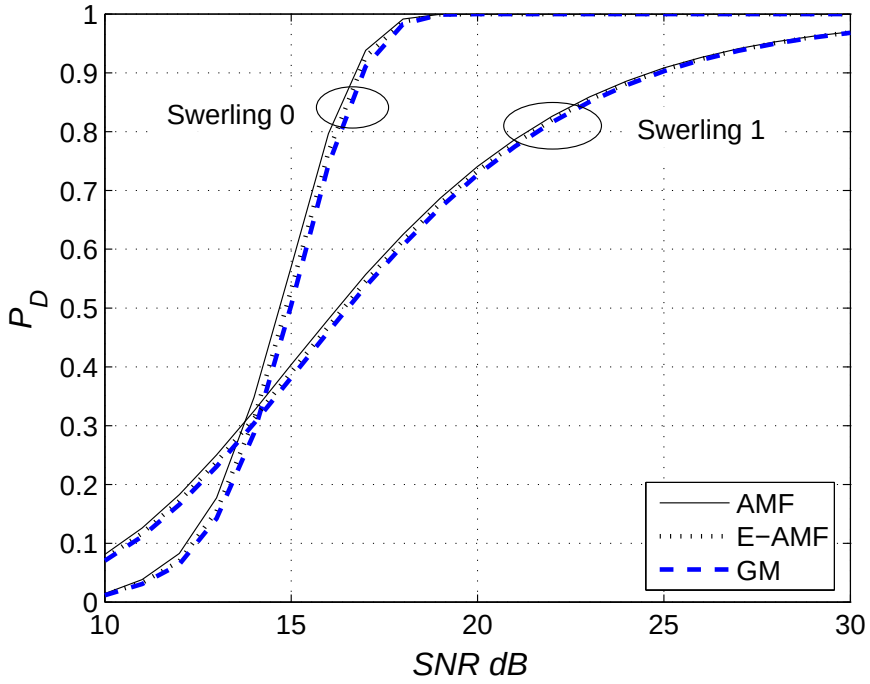

Fig. 6. Detection performances in homogeneous clutter. $L = 128$, $N = 64$, and $\alpha_o = 10^{-6}$.

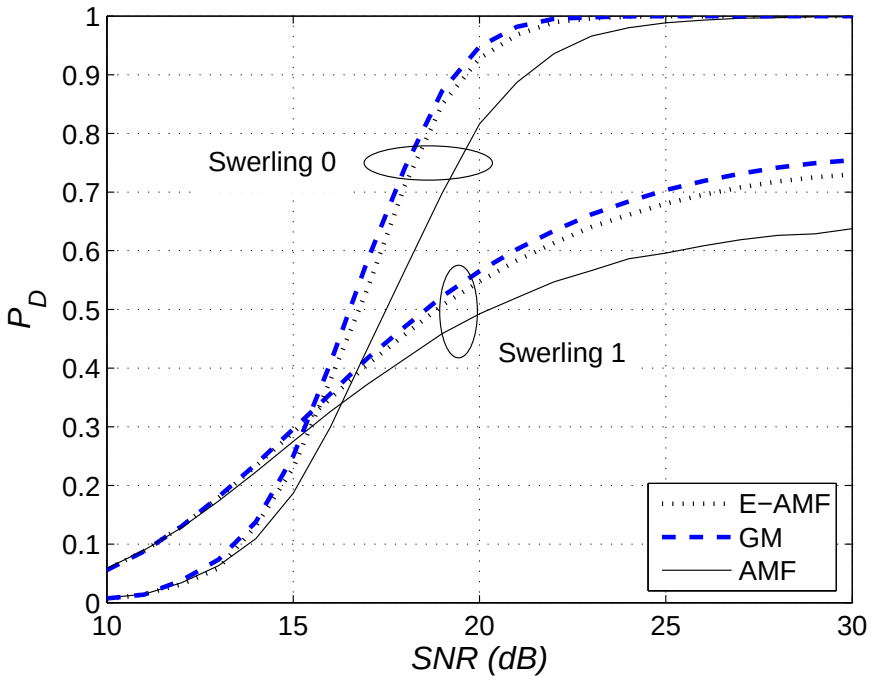

Fig. 7. Detection with 2 interfering targets in training data. $L = 128$, $N = 64$, and $\alpha_o = 10^{-6}$.